

GenCore version 5.1.4.P5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 17, 2003, 11:35:58 ; Search time 21.6471 Seconds
(Without alignments)
62.524 Million cell updates/sec

Title: US-09-855-754B-8

Perfect score: 245
Sequence: 1 ORATIRGDAPAGAVPGCA.....PLDGMVGVDSSTVDLAQ 46

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCTUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	245	100.0	911	4	US-08-460-269C-4	Sequence 4, Appl
2	232.5	94.9	922	4	US-08-460-269C-6	Sequence 6, Appl
3	202	82.4	910	4	US-08-460-269C-2	Sequence 2, Appl
4	74.5	30.4	745	2	US-09-010-928B-28	Sequence 28, Appl
5	74.5	30.4	870	2	US-09-010-928B-2	Sequence 2, Appl
6	70	28.6	907	2	US-09-010-928B-4	Sequence 4, Appl
7	63	25.7	20	2	US-09-010-928B-9	Sequence 9, Appl
8	59	24.1	20	2	US-09-010-928B-7	Sequence 7, Appl
9	59	24.1	420	4	US-08-845-998-8	Sequence 8, Appl
10	59	24.1	420	4	US-09-206-537-8	Sequence 8, Appl
11	59	24.1	420	4	US-09-430-854-8	Sequence 8, Appl
12	59	24.1	472	3	US-08-348-518C-5	Sequence 5, Appl
13	59	24.1	472	3	US-08-476-509B-5	Sequence 5, Appl
14	59	24.1	1711	3	US-08-342-930-2	Sequence 2, Appl
15	58	23.7	269	5	PCT-US94-01149-20	Sequence 20, Appl
16	58	23.7	304	5	PCT-US94-01149-32	Sequence 32, Appl
17	57.5	23.5	1304	5	PCT-US94-01149-28	Sequence 28, Appl
18	57.5	23.5	222	5	PCT-US94-01149-6	Sequence 6, Appl
19	57.5	23.5	307	5	PCT-US94-01149-16	Sequence 16, Appl
20	57.5	23.5	566	4	US-09-232-468A-14	Sequence 14, Appl
21	57.5	23.5	566	4	US-08-686-968C-227	Sequence 227, App
22	57.5	23.5	572	2	US-08-453-848-9	Sequence 9, Appl
23	57.5	23.5	572	4	US-09-169-027-9	Sequence 9, Appl
24	57	23.3	4545	2	US-08-804-227C-14	Sequence 14, Appl
25	57	23.3	4550	2	US-08-804-227C-8	Sequence 8, Appl
26	57	23.3	4550	2	US-08-804-198-2	Sequence 2, Appl
27	56.5	23.1	348	1	US-08-229-781-50	Sequence 50, Appl

28	56.5	23.1	348	1	US-08-630-918-50	Sequence 50, Appl
29	56.5	23.1	348	4	US-09-004-422-50	Sequence 50, Appl
30	56.5	23.1	574	4	US-09-276-400-6	Sequence 6, Appl
31	56.5	23.1	574	4	US-09-448-076-6	Sequence 6, Appl
32	56.5	23.1	574	4	US-09-702-572-6	Sequence 6, Appl
33	56	22.9	119	4	US-08-556-978B-60	Sequence 60, Appl
34	56	22.9	119	4	US-09-247-806-9	Sequence 9, Appl
35	56	22.9	426	1	US-07-918-023-2	Sequence 2, Appl
36	56	22.9	566	4	US-09-232-468A-22	Sequence 22, Appl
37	56	22.9	714	4	US-08-556-978B-61	Sequence 61, Appl
38	56	22.9	714	4	US-09-247-806-10	Sequence 10, Appl
39	54	22.0	595	2	US-08-425-069-4	Sequence 4, Appl
40	54	22.0	595	2	US-08-317-844B-4	Sequence 4, Appl
41	53.5	21.8	493	1	US-08-556-978B-59	Sequence 59, Appl
42	53.5	21.8	737	1	US-08-188-582-16	Sequence 16, Appl
43	53.5	21.8	737	1	US-08-646-715-16	Sequence 16, Appl
44	53	21.6	239	4	US-09-372-422A-42	Sequence 42, Appl
45	53	21.6	529	4	US-09-247-806-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-460-269C-4
Sequence 4, Application US/08460269C
Patent No. 6197548
GENERAL INFORMATION:
APPLICANT: CLARE, JEFFREY J.
ROMANOS, MICHAEL A.
TITLE OF INVENTION: EXPRESSION OF HETEROLOGOUS PROTEIN IN YEAST
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Millen, White, Zelano & Branigan, P.C.
STREET: 2200 Clarendon Blvd., Suite 1400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22201
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,269C
FILING DATE: 02-Jun-1995
ATTORNEY/AGENT INFORMATION:
NAME: Lebovitz, Richard M.
REGISTRATION NUMBER: 37,067
REFERENCE/DOCKET NUMBER: Popov-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 243-6333
TELEFAX: (703) 243-6410
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 911 amino acids
TYPE: amino acid
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-08-460-269C-4

Query Match 100.0% ; Score 245 ; DB 4 ; Length 911 ;
Best Local Similarity 100.0% ; Pred. No. 3.4e-23 ;
Matches 46 ; Conservative 0 ; Mismatches 0 ; Indels 0 ; Gaps 0 ;

QY 1 ORATIRGDAPAGAVPGGFGPLDGMVGVDSSTVDLAQ 46
DB 254 ORATIRGDAPAGAVPGGFGPLDGMVGVDSSTVDLAQ 299

```

? APPLICATION NUMBER: US/08/460,269C
? FILING DATE: 02-Jun-1995
? ATTORNEY/AGENT INFORMATION:
? NAME: Lepovitz, Richard M.
? REGISTRATION NUMBER: 37,067
? REFERENCE/DOCKET NUMBER: Popov-2
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (703) 243-6333
? TELEFAX: (703) 243-6410
? INFORMATION FOR SEQ ID NO: 2:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 910 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-08-460-269C-2

Query Match      82.4%; Score 202; DB 4; Length 910;
Best Local Similarity 73.2%; Pred. No. 1e-17; 2; Indels 10; Gaps 1
Matches 41; Conservative 3; Mismatches 2; Indels 10; Gaps 1

QY 1 QARTIRGAPAGAVPGAV-----PGGFPLLDGWYGVSDSYDLAQ 46
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 254 QGRATIRGDPALAGAVPGAVPGAGVPGGFGPGGFPVLDDGWYGVSDSYVELAQ 309

RESULT 4
US-09-010-9228B-28
? Sequence 28, Application US/09010928B
? Patent No. 5994099
? GENERAL INFORMATION:
? APPLICANT: Lewis, Randolph V
? APPLICANT: Hayashi, Cheryl Y
? TITLE OF INVENTION: EXTREMELY ELASTIC SPIDER SILK PROTEIN AND DNA
? TITLE OF INVENTION: CODING THEREFOR
? NUMBER OF SEQUENCES: 29
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
? STREET: 8110 GATEHOUSE RD. SUITE 500E
? CITY: FALLS CHURCH
? STATE: VIRGINIA
? COUNTRY: UNITED STATES OF AMERICA
? ZIP: 22042
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/010,928B
? FILING DATE: 22-JAN-1998
? CLASSIFICATION: 435
? ATTORNEY/AGENT INFORMATION:
? NAME: Murphy Jr., Gerald M
? REGISTRATION NUMBER: 28977
? REFERENCE/DOCKET NUMBER: 1447-109P
? INFORMATION FOR SEQ ID NO: 28:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 745 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
US-09-010-928B-28

Query Match      30.4%; Score 74.5; DB 2; Length 745;
Best Local Similarity 63.0%; Pred. No. 0.14; 7; Indels 3; Gaps 1;
Matches 17; Conservative 0; Mismatches 7; Indels 3; Gaps 1;
Db 248 GGAGPGAGPGGAGPGGAGP---GGYG 271

```

```

1 CLASSIFICATION: 435
2 ATTORNEY/AGENT INFORMATION:
3 NAME: Murphy Jr., Gerald M
4 REGISTRATION NUMBER: 28977
5 REFERENCE/DOCKET NUMBER: 1447-109P
6 INFORMATION FOR SEQ ID NO: 4:
7 SEQUENCE CHARACTERISTICS:
8 LENGTH: 907 amino acids
9 TYPE: amino acid
10 TOPOLOGY: linear
11 MOLECULE TYPE: protein
12 US-09-010-928B-4
13
14 Query Match 28.6% Score 70; DB 2; Length 907;
15 Best Local Similarity 65.0%; Pred. No. 0.68;
16 Matches 13; Conservative 1; Mismatches 6; Indels 0; Gaps 0.
17
18 QY 8 GDAPAGGAVPGGAVPGGFGP 27
19 | : ||| ||| |||||
20 Db 143 GSGPGGAGPGGAVPGGFGP 162
21
22 RESULT 7
23 US-09-010-928B-9
24 Sequence 9, Application US/09010928B
25 Patent No. 5994099
26 GENERAL INFORMATION:
27 APPLICANT: Lewis, Randolph v
28 APPLICANT: Hayashi, Cheryl Y
29 TITLE OF INVENTION: EXTREMELY ELASTIC SPIDER SILK PROTEIN AND DNA
30 TITLE OF INVENTION: CODING THEREFOR
31 NUMBER OF SEQUENCES: 29
32 CORRESPONDENCE ADDRESS:
33 ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
34 STREET: 8110 GATEHOUSE RD. SUITE 500E
35 CITY: FALLS CHURCH
36 STATE: VIRGINIA
37 COUNTRY: UNITED STATES OF AMERICA
38 ZIP: 22042
39 COMPUTER READABLE FORM:
40 MEDIUM TYPE: floppy disk
41 COMPUTER: IBM PC compatible
42 OPERATING SYSTEM: PC-DOS/MS-DOS
43 SOFTWARE: PatentIn Release #1.0, Version #1.30
44 CURRENT APPLICATION DATA:
45 APPLICATION NUMBER: US/09/010,928B
46 FILING DATE: 22-JAN-1998
47 CLASSIFICATION: 435
48 ATTORNEY/AGENT INFORMATION:
49 NAME: Murphy Jr., Gerald M
50 REGISTRATION NUMBER: 28977
51 REFERENCE/DOCKET NUMBER: 1447-109P
52 INFORMATION FOR SEQ ID NO: 9:
53 SEQUENCE CHARACTERISTICS:
54 LENGTH: 20 amino acids
55 TYPE: amino acid
56 TOPOLOGY: linear
57 MOLECULE TYPE: protein
58 US-09-010-928B-9
59
60 Query Match 25.7% Score 63; DB 2; Length 20;
61 Best Local Similarity 80.0%; Pred. No. 0.068;
62 Matches 12; Conservative 0; Mismatches 3; Indels 0; Gaps 0.
63
64 QY 13 GGAVPGGAVPGGFGP 27
65 ||| ||| ||| |||
66 Db 2 GGAGPGGAGPGGAGP 16
67
68 RESULT 8
69 US-09-010-928B-7
70 Sequence 7, Application US/09010928B
71 Patent No. 5994099

```

GENERAL INFORMATION:
APPLICANT: Lewis, Randolph V
TITLE OF INVENTION: EXTREMELY ELASTIC SPIDER SILK PROTEIN AND DNA
TITLE OF INVENTION: CODING THEREFOR
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
STREET: 8110 GATEHOUSE RD, SUITE 500E
CITY: FALLS CHURCH
STATE: VIRGINIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 22042
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/010,928B
FILING DATE: 22-JAN-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Murphy Jr., Gerald M
REGISTRATION NUMBER: 28977
REFERENCE/DOCKET NUMBER: 1447-109P
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-010-928B-7

Query Match 24.1%; Score 59; DB 2; Length 20;
Best Local Similarity 66.7%; Pred. No. 0.22;
Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 13 GGAVPGGAVPGGFGP 27
DB 2 GGYGPGGGGPGGYP 16

RESULT 9
US-08-845-998-8
Sequence 8, Application US/08845998
Patent No. 5879892
GENERAL INFORMATION:
APPLICANT: Van Baren, Nicolas
APPLICANT: Coulle, Pierre G.
APPLICANT: De Smet, Charles
APPLICANT: Lucas, Sophie
APPLICANT: Boon, Thierry
TITLE OF INVENTION: LEUKEMIA ASSOCIATED GENES
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
STREET: 600 Atlantic Avenue
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/845,998
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Van Amsterdam, John R.

REGISTRATION NUMBER: 40,212
REFERENCE/DOCKET NUMBER: L0461/7008
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)720-3500
TELEFAX: (617)720-2441
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 420 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-845-998-8

Query Match 24.1%; Score 59; DB 2; Length 420;
Best Local Similarity 57.9%; Pred. No. 7.1;
Matches 11; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 8 GDAPAGGAVPGGAVPGGFG 26
DB 150 GGGPGGGGGGPGGGGPGGGG 168

RESULT 10
US-09-206-537-8
Sequence 8, Application US/09206537
Patent No. 6130052
GENERAL INFORMATION:
APPLICANT: Van Baren, Nicolas
APPLICANT: Coulle, Pierre G.
APPLICANT: De Smet, Charles
APPLICANT: Lucas, Sophie
APPLICANT: Boon, Thierry
TITLE OF INVENTION: LEUKEMIA ASSOCIATED GENES
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
STREET: 600 Atlantic Avenue
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/206,537
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/845,998
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Van Amsterdam, John R.
REGISTRATION NUMBER: 40,212
REFERENCE/DOCKET NUMBER: L0461/7008
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)720-3500
TELEFAX: (617)720-2441
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 420 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-206-537-8

Query Match 24.1%; Score 59; DB 4; Length 420;
Best Local Similarity 57.9%; Pred. No. 7.1;
Matches 11; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 8 GDAPAGGAVPGGAVPGGFG 26

Db 150 GCGGCGGGGPGCGGPGGCG 168

RESULT 11

US-09-430-854-8
; Sequence 8, Application US/09430854
; Patent No. 6271019

GENERAL INFORMATION:

APPLICANT: Van Baren, Nicolas
APPLICANT: Coultre, Pierre G.
APPLICANT: De Smet, Charles

APPLICANT: Lucas, Sophie
APPLICANT: Boon, Thierry

TITLE OF INVENTION: LEUKEMIA ASSOCIATED GENES

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSEE: Wolf, Greenfield & Sacks, P.C.

STREET: 600 Atlantic Avenue

CITY: Boston

STATE: MA

COUNTRY: US

ZIP: 02210

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/430,854

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/845,998

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Van Amsterdam, John R.

REGISTRATION NUMBER: 40,212

REFERENCE/DOCKET NUMBER: L0461/7008

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617)720-3500

TELEFAX: (617)720-2441

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 420 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-430-854-8

Query Match 24.1%; Score 59; DB 4; Length 420;

Best Local Similarity 57.9%; Pred. No. 7.1;

Matches 11; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Y 8 GDAPGAVPGGAVPGGFG 26

Db 150 GCGGCGGGGPGCGGPGGCG 168

RESULT 12

US-08-348-518C-5
; Sequence 5, Application US/08348518C
; Patent No. 6022740

GENERAL INFORMATION:

APPLICANT: SUDOL, MARIUS

APPLICANT: PEER, BORK

APPLICANT: HENRY, CHEN

TITLE OF INVENTION: A SH2 DOMAIN ASSOCIATED PROTEIN, A

TITLE OF INVENTION: SIGNALING DOMAIN THEREIN, NUCLEIC ACIDS ENCODING THE

TITLE OF INVENTION: PROTEIN AND THE DOMAIN, AND DIAGNOSTIC AND THERAPEUTIC USES

NUMBER OF SEQUENCES: 26

CORRESPONDENCE ADDRESS:

ADDRESSEE: Klauder & Jackson

STREET: 411 Hackensack Avenue

CITY: Hackensack

STATE: New Jersey

COUNTRY: USA

ZIP: 07601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/348,518C

FILING DATE: 01-DEC-1994

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: Jackson Esq., David A.

REGISTRATION NUMBER: 26,742

REFERENCE/DOCKET NUMBER: 600-1-101

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201 487-5800

TELEFAX: 201 343-1684

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 472 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ORGANISM: mouse

IMMEDIATE SOURCE:

CLONE: YAP

US-08-348-518C-5

Query Match 24.1%; Score 59; DB 3; Length 472;

Best Local Similarity 42.4%; Pred. No. 8.1;

Matches 14; Conservative 1; Mismatches 18; Indels 0; Gaps 0;

Y 10 APAGAVPGGAVPGGFGPLDGMVGVSDSTV 42

Db 199 APAGAVPGGAVPGGFGPLDGMVGVSDSTV 231

RESULT 13

US-08-476-509B-5
; Sequence 5, Application US/08476509B
; Patent No. 6034212

GENERAL INFORMATION:

APPLICANT: SUDOL, MARIUS

APPLICANT: PEER, BORK

APPLICANT: HENRY, CHEN

TITLE OF INVENTION: A SH2 DOMAIN ASSOCIATED PROTEIN, A

TITLE OF INVENTION: SIGNALING DOMAIN THEREIN, NUCLEIC ACIDS ENCODING THE

TITLE OF INVENTION: PROTEIN AND THE DOMAIN, AND DIAGNOSTIC AND THERAPEUTIC USES

NUMBER OF SEQUENCES: 50

CORRESPONDENCE ADDRESS:

ADDRESSEE: Klauder & Jackson

STREET: 411 Hackensack Avenue

CITY: Hackensack

STATE: New Jersey

COUNTRY: USA

ZIP: 07601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/476,509B

FILING DATE: 01-DEC-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-101 CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 472 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: NO
ORIGINAL SOURCE:
ORGANISM: Mouse
IMMEDIATE SOURCE:
CLONE: YAP
US-08-476-509B-5

Query Match 24.1%; Score 59; DB 3; Length 472;
Best Local Similarity 42.4%; Pred. No. 8.1;
Matches 14; Conservative 1; Mismatches 18; Indels 0; Gaps 0;

QY 10 APAGAVPGGAVPGGFGPLDGMVGVDSSTV 42
DB 199 APASPAVPORTLMSASGSLPFDGMEQANTODGEV 231

RESULT 14
US-08-342-930-2
Sequence 2, Application US/08342930
Patent No. 5821084
GENERAL INFORMATION:
APPLICANT: OLMSTED, ELIZABETH A.
APPLICANT: MAURO, LAURA J.
APPLICANT: DAVIS, ALAN R.
APPLICANT: DIXON, JACK E.
TITLE OF INVENTION: OSTEOBLAST-TESTICULAR PROTEIN TYROSINE
TITLE OF INVENTION: PHOSPHATASE
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 755 Page Mill Road
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/342,930
FILING DATE: 21-NOV-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: KONSKE, ANTOINETTE F.
REGISTRATION NUMBER: 34,202
REFERENCE/DOCKET NUMBER: 20344-20975.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 813-5600
TELEFAX: (415) 494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1711 amino acids
TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-342-930-2

Query Match 24.1%; Score 59; DB 2; Length 1711;
Best Local Similarity 38.0%; Pred. No. 35;
Matches 19; Conservative 8; Mismatches 19; Indels 4; Gaps 2;

QY 1 QRATIRRGADP---AGGAVPGGAVPGGFGPLDGMVGVDSSTVDLAQ 46
DB 416 RRALLYSDDAPGSLGNISVPSGATHVIFCGIYVGAHYRVDIASSTGDISQ 465

RESULT 15
PCT-US94-01149-20
Sequence 20, Application PC/TUS9401149
GENERAL INFORMATION:
APPLICANT: Shatzman, Allan
APPLICANT: Scott, Miller
APPLICANT: Dillon, Susan B.
APPLICANT: Kane, James
TITLE OF INVENTION: Vaccinal Polypeptides
NUMBER OF SEQUENCES: 72
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation - Corporate
STREET: Patents
CITY: U.S. Mailcode UM2220 - 709 Swedeland Road
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19406-2799
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/01149
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 149,150
FILING DATE: 05-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 013,415
FILING DATE: 01-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 108,914
FILING DATE: 18-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 837,773
FILING DATE: 18-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 751,896
FILING DATE: 30-AUG-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 387,200
FILING DATE: 28-JUL-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 238,801
FILING DATE: 02-NOV-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 645,732
FILING DATE: 30-AUG-1984
ATTORNEY/AGENT INFORMATION:
NAME: Baumeister, Kirk
REGISTRATION NUMBER: 33,833
REFERENCE/DOCKET NUMBER: P50134 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-270-5096
TELEFAX: 215-270-5090
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:

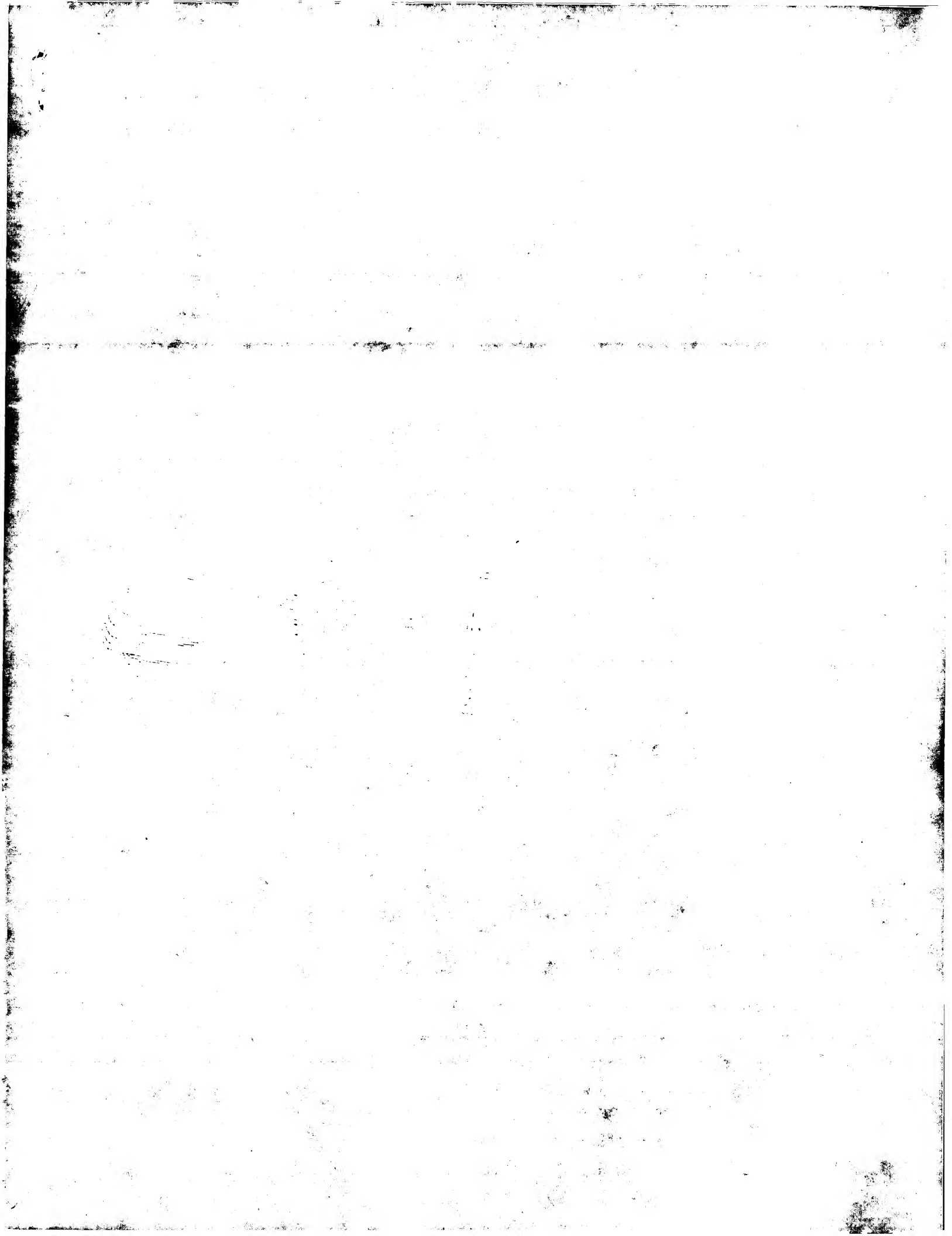
Mon Mar 17 12:14:30 2003

LENGTH: 269 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 PCT-US94-01149-20

Query Match 23.7%; Score 58; DB 5; Length 269;
 Best Local Similarity 31.8%; Pred. No. 5.7;
 Matches 14; Conservative 5; Mismatches 7; Indels 18; Gaps 2;

OY 8 GDAP-----AGGAVPGAVPGGEGPLLDGWTG 34
 |||||
 Db 28 GDAPFLDLRLRRDQKSMDSLRLFGAI-AGFIEGGWTGMIDGWTG 70

Search completed: March 17, 2003, 11:39:41
 Job time : 23.6471 secs



GenCore version 5.1.4.P5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 17, 2003, 11:32:04 ; Search time 60.4314 Seconds
(without alignments)
101.430 Million cell updates/sec

Title: US-09-855-754b-8
Perfect score: 245
Sequence: 1 QRATIRRGDAPAGGAVPGGA.....PLDGYGVDSSTVDLAQ 46

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	245	100.0	46	23 AAE16187	B. bronchiseptica
2	245	100.0	911	13 AAR14320	Pertactin antigen
3	245	100.0	911	13 AAR26503	Prn proteins. Bor
4	245	100.0	911	23 AAE16183	Bordetella bronchi
5	232.5	94.9	51	23 AAE16186	B. bronchiseptica
6	232.5	94.9	922	13 AAR14321	Pertactin antigen
7	232.5	94.9	922	13 AAR25578	Bordetella parapert
8	232.5	94.9	922	23 AAE16185	Bordetella pertuss
9	214.5	87.6	51	23 AAE16192	B. bronchiseptica
10	212	86.5	56	23 AAE16188	B. bronchiseptica

11	212	86.5	56	23 AAE16189	B. bronchiseptica
12	212	86.5	56	23 AAE16191	B. bronchiseptica
13	212	86.5	910	23 AAE16184	Bordetella pertuss
14	212	86.5	910	23 AAE17146	Bordetella pertuss
15	209.5	85.5	61	23 AAE16190	B. bronchiseptica
16	175	71.4	45	23 AAE17153	F647 monoclonal an
17	175	71.4	45	23 AAE17154	F647 monoclonal an
18	175	71.4	45	23 AAE17155	Pem4 monoclonal an
19	175	71.4	45	23 AAE17156	Pem6 monoclonal an
20	175	71.4	45	23 AAE17157	Pem7 monoclonal a
21	175	71.4	45	23 AAE17158	Pem7 monoclonal a
22	175	71.4	45	23 AAE17159	Bordetella pertuss
23	133.5	54.5	54	23 AAE17150	Bordetella pertuss
24	133.5	54.5	59	23 AAE17149	Bordetella pertuss
25	133.5	54.5	64	23 AAE17148	Bordetella pertuss
26	132.5	54.1	31	23 AAE17173	Maltose binding pr
27	132.5	54.1	36	23 AAE17172	Maltose binding pr
28	132.5	54.1	41	23 AAE17171	Maltose binding pr
29	128.5	52.4	36	23 AAE17174	Maltose binding pr
30	122	49.8	59	23 AAE17147	Bordetella pertuss
31	121	49.4	36	23 AAE17141	Bordetella pertuss
32	121	49.4	36	23 AAE17170	Maltose binding pr
33	120.5	49.2	54	23 AAE17151	Bordetella pertuss
34	120	49.0	56	23 AAE17152	Bordetella pertuss
35	86.5	35.3	20	23 AAE17144	Bordetella pertuss
36	86	35.1	15	23 AAE17166	Monoclonal Ab derl
37	85	34.7	15	23 AAE17163	Monoclonal Ab derl
38	81	33.1	15	23 AAE17161	Monoclonal Ab derl
39	77	31.4	15	23 AAE17160	Monoclonal Ab derl
40	73.5	30.0	438	21 AAG17460	Arabidopsis thalia
41	73.5	30.0	460	21 AAG17459	Arabidopsis thalia
42	73.5	30.0	646	21 AAG17458	Arabidopsis thalia
43	73	29.8	15	23 AAE17165	Monoclonal Ab derl
44	72	29.4	15	23 AAE17164	Bordetella pertuss
45	68	27.8	23	23 AAE17162	Monoclonal Ab derl

ALIGNMENTS

RESULT 1	AAE16187	standard; peptide; 46 AA.
ID	AAE16187	
AC	AAE16187;	
XX		
DT	26-MAR-2002	(first entry)
XX		
DE	B. bronchiseptica strain I-2 pertactin outer membrane protein region I.	
KW	Pertactin; PRN: outer membrane protein; vaccine; Bordetella infection;	
KW	therapy; antibiotic; antibacterial; region I.	
XX		
OS	Bordetella bronchiseptica.	
XX		
PN	W0200190143-A2.	
XX		
PD	29-NOV-2001.	
XX		
PR	23-MAY-2001; 2001WO-EP06457.	
XX		
PR	25-MAY-2000; 2000US-206969P.	
XX		
PA	(INST) INST PASTEUR.	
XX		
PI	Gulso-maclouf N, Boursaux-eude C;	
XX		
DR	WPI; 2002-097639/13.	
XX		
PT	Polypeptides containing polymorphisms of the repeated regions of	
PT	pertactin in Bordetella species, useful in immunogenic compositions for	
PT	treating infections caused by Bordetella and in diagnostic methods	
XX		

PS Claim 26; Fig 1b; 47pp; English.
 CC The present invention relates to Bordetella bronchiseptica pertactin
 CC (outer membrane protein) or their fragments. Pertactin (PRN) is used as
 CC vaccine. Pertactin antibody is useful for treating Bordetella infections
 CC and used to detect Bordetella antigens in biological preparations or in
 CC purifying corresponding proteins, glycoproteins or their mixtures when
 CC used in affinity chromatographic columns. Pertactin is useful as antigens
 CC to identify antibodies to Bordetella in materials such as human or other
 CC animal tissue and human or other animal body fluids, including human sera,
 CC fluids, such as human or other animal body fluids, including human sera,
 CC and to determine the concentration of Ab in those materials. Thus the
 CC antigens can be used for qualitative or quantitative determination of
 CC pertactin in a material. The present sequence is B. bronchiseptica
 CC pertactin outer membrane protein region I.
 SQ Sequence 46 AA;

Query Match 100.0%; Score 245; DB 23; Length 46;
 Best Local Similarity 100.0%; Pred. No. 2,9e-22;
 Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 QRATIRRGDAPAGGAVPGGFGPPLDGMVGVDVSDSTVDLAQ 46
 DB 1 QRATIRRGDAPAGGAVPGGFGPPLDGMVGVDVSDSTVDLAQ 46

RESULT 2

ID AAR14320 standard; Protein; 911 AA.

AC AAR14320;

DT 20-JAN-1992 (first entry)

DE Pertactin antigen P.68.

KW Pertactin; Pichia; B. pertussis; B. paraptussis.

OS Bordetella bronchiseptica.

FN Key Location/Qualifiers

FT Peptide 266..270 /label= repeat

FT Peptide 271..275 /label= repeat

FT Peptide 570..572 /label= repeat

FT Peptide 574..576 /label= repeat

FT Peptide 578..580 /label= repeat

FT Peptide 581..583 /label= repeat

FT Peptide 584..586 /label= repeat

FT Peptide 587..589 /label= repeat

FT Peptide 599..601 /label= repeat

FT Peptide /label= repeat

FN Peptide /label= repeat

FN Peptide /label= repeat

FN Peptide /label= repeat

FN Peptide /label= repeat

FN Peptide /label= repeat

FN Peptide /label= repeat

FN Peptide /label= repeat

FN Peptide /label= repeat

FN Peptide /label= repeat

FN Peptide /label= repeat

FN Peptide /label= repeat

FN Peptide /label= repeat

FN Peptide /label= repeat

FN Peptide /label= repeat

FN Peptide /label= repeat

FN Peptide /label= repeat

DR WPI: 1991-325214/44.

DR N-PSDB; AA014319.

XX Pichia microorganism transformants - for production of

PT Bordetella pertactin antigens for whooping cough vaccines

PS Disclosure; Fig 1b; 38pp; English.

XX Pichia microorganisms are transformed for the expression of

CC pertactin antigens. DNA sequence used are represented in AA014319-20

CC encoding the B. bronchiseptica P.68 and B. paraptussis P.70 antigen

CC respectively or the B. pertussis P.69 encoding sequence described

CC by I.G. Charles et al. Proc. Natl. Acad. Sci. USA, Vol. 80:3534-3448

CC (1989).

XX Sequence 911 AA;

Query Match 100.0%; Score 245; DB 12; Length 911;
 Best Local Similarity 100.0%; Pred. No. 5.8e-21;
 Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 QRATIRRGDAPAGGAVPGGFGPPLDGMVGVDVSDSTVDLAQ 46
 DB 254 QRATIRRGDAPAGGAVPGGFGPPLDGMVGVDVSDSTVDLAQ 299

RESULT 3

ID AAR26503 standard; Protein; 911 AA.

AC AAR26503;

DT 12-MAR-1993 (first entry)

DE prn proteins.

KW B. bronchiseptica; P.68; outer membrane protein; piglet; probe;

OS atrophic rhinitis; alternative cleavage.

OS Bordetella bronchiseptica.

FN Key Location/Qualifiers

FT Protein 35..632 /label= P.68

FT Region 266..279 /label= Repeat_region

FT Region 570..589 /label= Repeat_region

FT Region 574..586 /label= Repeat_region

FT Peptide 260..262 /label= RGD_tripeptide

FT Peptide 701..703 /label= RGD_tripeptide

FT Peptide /label= RGD_tripeptide

FN Peptide /label= RGD_tripeptide

FN Peptide /label= RGD_tripeptide

FN Peptide /label= RGD_tripeptide

FN Peptide /label= RGD_tripeptide

FN Peptide /label= RGD_tripeptide

FN Peptide /label= RGD_tripeptide

FN Peptide /label= RGD_tripeptide

FN Peptide /label= RGD_tripeptide

FN Peptide /label= RGD_tripeptide

FN Peptide /label= RGD_tripeptide

FN Peptide /label= RGD_tripeptide

FN Peptide /label= RGD_tripeptide

FN Peptide /label= RGD_tripeptide

FN Peptide /label= RGD_tripeptide

FN Peptide /label= RGD_tripeptide

FN Peptide /label= RGD_tripeptide

FN Peptide /label= RGD_tripeptide

FN Peptide /label= RGD_tripeptide

DNA encoding a Bordetella bronchiseptica protein - used for
 obtaining vaccines for preventing respiratory diseases, partic.
 atrophic rhinitis in pigs

Claim 1; Fig 1; 28pp; English.

XX The sequence given is the P.94 antigen from B. bronchiseptica. The
 CC P.68 antigen is formed by alternative cleavage of this protein.
 CC P.68 is an outer membrane protein with a molecular weight of 68 kD
 CC which is associated with protection of piglets against atrophic
 CC rhinitis. The DNA sequence encoding these proteins was derived by
 CC standard recombinant DNA techniques using P.68 probes to isolate the
 CC entire P.94 sequence.
 XX
 SO Sequence 911 AA:
 Query Match 100.0%; Score 245; DB 13; Length 911;
 Best Local Similarity 100.0%; Pred. No. 5.8e-21;
 Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QRATIRRGDAPAGAVPGGAVPGFPLLDGWTGVDSSTVDLAQ 46
 DB 254 QRATIRRGDAPAGAVPGGAVPGGAVPGFPLLDGWTGVDSSTVDLAQ 299
 RESULT 4
 ID AAE16183 standard; Protein; 911 AA.
 AC AAE16183;
 DE 26-MAR-2002 (first entry)
 XX
 DE Bordetella bronchiseptica pertactin outer membrane protein, p.68.
 XX
 KW Pertactin; PRN; outer membrane protein; vaccine; Bordetella infection;
 KM therapy; antibiotic; antibacterial; p.68.
 XX
 OS Bordetella bronchiseptica.
 XX
 FH Key Location/Qualifiers
 FT Region 254..299
 FT /note= "Pertactin region I"
 FT Region 559..610
 FT /note= "Pertactin region II"
 XX
 PN WO200190143-A2.
 XX
 PD 29-NOV-2001.
 XX
 PF 23-MAY-2001; 2001WO-EP06457.
 XX
 PR 25-MAY-2000; 2000US-206969P.
 XX
 PA (INSP) INST PASTEUR.
 PI Guiso-maclouf N, Boursaux-eude C;
 XX
 DR WPI: 2002-097639/13.
 DR N-PSDB: AAD26440.
 XX
 PT Polypeptides containing polymorphisms of the repeated regions of
 PT pertactin in Bordetella species, useful in immunogenic compositions for
 PT treating infections caused by Bordetella and in diagnostic methods -
 XX
 PS Disclosure; Page 28; 47pp; English.
 XX
 CC The present invention relates to Bordetella bronchiseptica pertactin
 CC (outer membrane protein) or their fragments. Pertactin (PRN) is used as
 CC vaccine. Pertactin antibody is useful for treating Bordetella infections
 CC and used to detect Bordetella antigens in biological preparations or in
 CC purifying corresponding proteins, glycoproteins or their mixtures when
 CC used in affinity chromatographic columns. Pertactin is useful as antigens
 CC to identify antibodies to Bordetella in materials such as human or other
 CC animal tissue and human or other animal cells, as well as biological
 CC fluids, such as human or other animal body fluids, including human sera,
 CC and to determine the concentration of Ab in those materials. Thus the
 CC antigens can be used for qualitative or quantitative determination of

CC Bordetella in a material. The present sequence is B. bronchiseptica
 CC pertactin outer membrane protein, p.68.
 XX
 SO Sequence 911 AA:
 Query Match 100.0%; Score 245; DB 23; Length 911;
 Best Local Similarity 100.0%; Pred. No. 5.8e-21;
 Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QRATIRRGDAPAGAVPGGAVPGFPLLDGWTGVDSSTVDLAQ 46
 DB 254 QRATIRRGDAPAGAVPGGAVPGGAVPGFPLLDGWTGVDSSTVDLAQ 299
 RESULT 5
 ID AAE16186 standard; peptide; 51 AA.
 AC AAE16186;
 DE 26-MAR-2002 (first entry)
 XX
 DE B. bronchiseptica strain I-1 pertactin outer membrane protein region I.
 XX
 KW Pertactin; PRN; outer membrane protein; vaccine; Bordetella infection;
 KM therapy; antibiotic; antibacterial; region I.
 XX
 OS Bordetella bronchiseptica.
 XX
 PN WO200190143-A2.
 XX
 PD 29-NOV-2001.
 XX
 PF 23-MAY-2001; 2001WO-EP06457.
 XX
 PR 25-MAY-2000; 2000US-206969P.
 XX
 PA (INSP) INST PASTEUR.
 PI Guiso-maclouf N, Boursaux-eude C;
 XX
 DR WPI: 2002-097639/13.
 XX
 PD Polypeptides containing polymorphisms of the repeated regions of
 PD pertactin in Bordetella species, useful in immunogenic compositions for
 PD treating infections caused by Bordetella and in diagnostic methods -
 XX
 PS Claim 26; Fig 1b; 47pp; English.
 XX
 CC The present invention relates to Bordetella bronchiseptica pertactin
 CC (outer membrane protein) or their fragments. Pertactin (PRN) is used as
 CC vaccine. Pertactin antibody is useful for treating Bordetella infections
 CC and used to detect Bordetella antigens in biological preparations or in
 CC purifying corresponding proteins, glycoproteins or their mixtures when
 CC used in affinity chromatographic columns. Pertactin is useful as antigens
 CC to identify antibodies to Bordetella in materials such as human or other
 CC animal tissue and human or other animal cells, as well as biological
 CC fluids, such as human or other animal body fluids, including human sera,
 CC and to determine the concentration of Ab in those materials. Thus the
 CC antigens can be used for qualitative or quantitative determination of
 CC Bordetella in a material. The present sequence is B. bronchiseptica
 CC pertactin outer membrane protein region I.
 XX
 SO Sequence 51 AA:
 Query Match 94.9%; Score 232.5; DB 23; Length 51;
 Best Local Similarity 90.2%; Pred. No. 1e-20;
 Matches 46; Conservative 0; Mismatches 0; Indels 5; Gaps 1;
 QY 1 QRATIRRGDAPAGAVPGGAVPGFPLLDGWTGVDSSTVDLAQ 46
 DB 1 QRATIRRGDAPAGAVPGGAVPGGAVPGFPLLDGWTGVDSSTVDLAQ 51

```

RESULT 6
AAR14321 standard; Protein: 922 AA.
ID AAR14321
XX
AC AAR14321;
XX
DT 20-JAN-1992 (first entry)
XX
DE Pertactin antigen P.70.
XX
KM Pertactin; Pichia; B. pertussis; B.bronchiseptica; ss.
XX
OS Bordetella parapertussis.
XX
FH Key Location/Qualifiers
FT Peptide 266..270 /label= repeat
FT Peptide 271..275 /label= repeat
FT Peptide 276..280 /label= repeat
FT Peptide 281..285 /label= repeat
FT Peptide 575..577 /label= repeat
FT Peptide 579..581 /label= repeat
FT Peptide 583..585 /label= repeat
FT Peptide 586..588 /label= repeat
FT Peptide 589..591 /label= repeat
FT Peptide 592..594 /label= repeat
FT Peptide 595..597 /label= repeat
FT Peptide 598..600 /label= repeat
FT Peptide 610..612 /label= repeat
FT Peptide /label= repeat
XX
PN WO9115571-A.
XX
PD 17-OCT-1991.
XX
PF 28-MAR-1991; 91WO-GB00487.
XX
PR 02-APR-1990; 90GB-0007416.
XX
PA (WELL ) WELLCOME FOUNDATION LTD.
XX
PI Clare JJ, Romanos MA;
XX
DR MPI: 1991-325214/44.
XX
DR N-PSDB; AAO14320.
XX
PT Bordetella parapertussis transformants - for production of
XX PT Bordetella parapertussis antigens for whooping cough vaccines
XX PS Disclosure; Fig 1C; 38pp; English.
XX
CC Pichia microorganisms are transformed for the expression of
CC pertactin antigens. DNA sequence used are represented in AAO14319-20
CC encoding the B. bronchiseptica P.68 and B. parapertussis P.70 antigen
CC respectively or the B. pertussis P.69 encoding sequence described
CC by I.G. Charles et al. Proc. Natl. Acad. Sci. USA, Vol. 80:3554-3448
CC (1989).
CC The 46 Xs represent amino acids missing in the specification.
CC
SQ Sequence 922 AA;

```

```

Query Match 94.9%; Score 232.5; DB 12; Length 922;
Best Local Similarity 90.2%; Pred. No. 1; ee-19;
Matches 46; Conservative 0; Mismatches 0; Indels 5; Gaps 1;
Dy 1 QARIRRGDAPAGAV-----PGGAVPGGFGPLLDGMYGVDSYVDLAQ 46
Db 254 QARIRRGDAPAGAVPGGXPGGAVPGGFGPLLDGMYGVDSYVDLAQ 304

```

```

RESULT 7
AAR25578 standard; Protein: 922 AA.
ID AAR25578
XX
AC AAR25578;
XX
DT 08-JAN-1993 (first entry)
XX
DE Bordetella parapertussis P95 antigen precursor.
XX
KM Whooping cough; P70 antigen; P95 precursor protein; vaccination.
XX
OS Bordetella parapertussis.
XX
FH Key Location/Qualifiers
FT Protein 35..643
FT Binding-site 260..262 /label= P70
FT Region 266..285 /note="motif associated with cell-cell adhesion"
FT Region 575..612 /note="contains 5 direct, tandem repeats"
FT Binding-site 712..714 /note="contains 9 direct repeats of Pro-Gln-Pro"
FT Binding-site /note="motif associated with cell-cell adhesion"
XX
PN WO9211292-A.
XX
PD 09-JUL-1992.
XX
PF 23-DEC-1991; 91WO-GB02302.
XX
PR 21-DEC-1990; 90GB-0027901.
XX
PA (WELL ) WELLCOME FOUND LTD.
XX
PI Charles IG;
XX
DR MPI: 1992-250033/30.
XX
DR N-PSDB; AAO26509.
XX
PT Acellular vaccine for immunisation against whooping cough -
XX PT comprises protein uncontaminated by B. parapertussis components
XX and capable of binding antibodies which bind native P70 antigen
XX
PS Claim 1; Fig 1; 20pp; English.
XX
CC A cosmid library was constructed by transforming E.coli HB101 with
CC recombinant cosmids prepared by partial digestion of B.parapertussis
CC chromosomal DNA with Sma3a and cloning of 40-50kb fragments into the
CC BamHI site of cosmid pHC79. The cosmids were screened with a 1.8kb
CC ClaI fragment from the prn gene of B.pertussis. The insert from one
CC positive colony, harbouring cosmid pBD811, was sequenced and found to
CC contain an open reading frame encoding a 922 amino acid protein
CC with calculated mol.wt. 95,177. This precursor protein ("P95") is
CC processed in vivo to the P70 antigen of apparent mol. wt. 70,000 as
CC determined by SDS-PAGE, but with actual mol.wt. 61kD. Antigenic
CC fragments of the protein will be useful in developing an acellular
CC vaccine against B.parapertussis. Preferred fragments include amino
CC acids Pro577 to Pro612 or Ala574 to Pro612.
CC
SQ Sequence 922 AA;

```

Query Match 94.9%; Score 232.5; DB 13; Length 922;

Best Local Similarity 90.2%; Pred. No. 1.8e-19;
Matches 46; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

QY 1 ORATIRRGDAPAGAVPGGFGPPLLDGMYGVDSSTVDLAQ 46
DB 254 ORATIRRGDAPAGAVPGGFGPPLLDGMYGVDSSTVDLAQ 304

RESULT 8

AAE16185 standard; Protein; 922 AA.

AC AAE16185;
DT 26-MAR-2002 (first entry)

DE Bordetella parapertussis pertactin outer membrane protein, p.70.

KW Pertactin; PRN; outer membrane protein; vaccine; Bordetella infection;
KM therapy; antibiotic; antibacterial; p.70.

OS Bordetella parapertussis.

PH Key Location/Qualifiers

FT Region 254..304

FT /note="Pertactin region I"

FT 564..621

FT /note="Pertactin region II"

PN MO200190143-A2.

PD 29-NOV-2001.

PF 23-MAY-2001; 2001MO-EP06457.

PR 25-MAY-2000; 2000US-206969P.

PS (INSP) INST PASTEUR.

PA Guiso-maclouf N, Boursaux-eude C;

PI WPI: 2002-097639/13.

DR N-PSDB; AAD6442.

XX Polypeptides containing polymorphisms of the repeated regions of
PT pertactin in Bordetella species; useful in immunogenic compositions for
PT treating infections caused by Bordetella and in diagnostic methods

PS Disclosure; Page 34; 47pp; English.

XX The present invention relates to Bordetella bronchiseptica pertactin
CC (outer membrane protein) or their fragments. Pertactin (PRN) is used as
CC vaccine. Pertactin antibody is useful for treating Bordetella infections
CC and used to detect Bordetella antigens in biological preparations or in
CC purifying corresponding proteins, glycoproteins or their mixtures when
CC used in affinity chromatographic columns. Pertactin is useful as antigens
CC to identify antibodies to Bordetella in materials such as human or other
CC animal tissue and human or other animal body fluids, including human sera,
CC fluids, such as human or other animal body fluids, including human sera,
CC and to determine the concentration of Ab in those materials. Thus the
CC antigens can be used for qualitative or quantitative determination of
CC Bordetella in a material. The present sequence is B. parapertussis
CC pertactin outer membrane protein, p.70.

XX Sequence 922 AA;

Query Match 94.9%; Score 232.5; DB 23; Length 922;
Best Local Similarity 90.2%; Pred. No. 1.8e-19;

Matches 46; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

QY 1 ORATIRRGDAPAGAVPGGFGPPLLDGMYGVDSSTVDLAQ 46
DB 254 ORATIRRGDAPAGAVPGGFGPPLLDGMYGVDSSTVDLAQ 304

RESULT 9

AAE16192 standard; peptide; 51 AA.

AC AAE16192;

DT 26-MAR-2002 (first entry)

DE B. bronchiseptica strain prn4 pertactin outer membrane protein region I.

KW Pertactin; PRN; outer membrane protein; vaccine; Bordetella infection;
KM therapy; antibiotic; antibacterial; region I.

OS Bordetella bronchiseptica.

PN MO200190143-A2.

PD 29-NOV-2001.

PF 23-MAY-2001; 2001MO-EP06457.

PR 25-MAY-2000; 2000US-206969P.

PS (INSP) INST PASTEUR.

PA Guiso-maclouf N, Boursaux-eude C;

DR WPI: 2002-097639/13.

XX Polypeptides containing polymorphisms of the repeated regions of
PT pertactin in Bordetella species; useful in immunogenic compositions for
PT treating infections caused by Bordetella and in diagnostic methods

PS Disclosure; Fig 1b; 47pp; English.

XX The present invention relates to Bordetella bronchiseptica pertactin
CC (outer membrane protein) or their fragments. Pertactin (PRN) is used as
CC vaccine. Pertactin antibody is useful for treating Bordetella infections
CC and used to detect Bordetella antigens in biological preparations or in
CC purifying corresponding proteins, glycoproteins or their mixtures when
CC used in affinity chromatographic columns. Pertactin is useful as antigens
CC to identify antibodies to Bordetella in materials such as human or other
CC animal tissue and human or other animal body fluids, including human sera,
CC fluids, such as human or other animal body fluids, including human sera,
CC and to determine the concentration of Ab in those materials. Thus the
CC antigens can be used for qualitative or quantitative determination of
CC Bordetella in a material. The present sequence is B. bronchiseptica
CC pertactin outer membrane protein region I.

XX Sequence 51 AA;

Query Match 87.6%; Score 214.5; DB 23; Length 51;
Best Local Similarity 82.4%; Pred. No. 1.4e-18;
Matches 42; Conservative 3; Mismatches 1; Indels 5; Gaps 1;

QY 1 ORATIRRGDAPAGAVPGGFGPPLLDGMYGVDSSTVDLAQ 46
DB 1 ORATIRRGDAPAGAVPGGFGPPLLDGMYGVDSSTVDLAQ 51

RESULT 10

AAE16188 standard; peptide; 56 AA.

AC AAE16188;

DT 26-MAR-2002 (first entry)

DE B. bronchiseptica strain I-3 pertactin outer membrane protein region I.
KM Pertactin; PRN; outer membrane protein; vaccine; Bordetella infection;

KW therapy; antibiotic; antibacterial; region I.
 XX
 OS Bordetella bronchiseptica.
 XX
 PN WO200190143-A2.
 XX
 PD 29-NOV-2001.
 XX
 PF 23-MAY-2001; 2001WO-EP06457.
 XX
 PR 25-MAY-2000; 2000US-206969P.
 XX
 PA (INSP) INST PASTEUR.
 PI Guiso-maclouf N, Boursaux-eude C;
 DR WPI; 2002-097639/13.
 XX
 PT Polypeptides containing polymorphisms of the repeated regions of
 PT pertactin in Bordetella species, useful in immunogenic compositions for
 PT treating infections caused by Bordetella and in diagnostic methods -
 XX
 PS Claim 26; Fig 1b; 47pp; English.
 CC The present invention relates to Bordetella bronchiseptica pertactin
 CC (outer membrane protein) or their fragments. Pertactin (PRN) is used as
 CC vaccine. Pertactin antibody is useful for treating Bordetella infections
 CC and used to detect Bordetella antigens in biological preparations or in
 CC purifying corresponding proteins, glycoproteins or their mixtures when
 CC used in affinity chromatographic columns. Pertactin is useful as antigens
 CC to identify antibodies to Bordetella in materials such as human or other
 CC animal tissue and human or other animal cells, as well as biological
 CC fluids, such as human or other animal body fluids, including human sera,
 CC and to determine the concentration of Ab in those materials. Thus the
 CC antigens can be used for qualitative or quantitative determination of
 CC Bordetella in a material. The present sequence is B. bronchiseptica
 CC pertactin outer membrane protein region I.
 CC
 SQ Sequence 56 AA;
 Query Match 86.5%; Score 212; DB 23; Length 56;
 Best Local Similarity 75.0%; Pred. No. 3,1e-18;
 Matches 42; Conservative 2; Mismatches 2; Indels 10; Gaps 1;
 QY 1 ORATTIRGDAPAGGAVPGGAV-----PGGFGPLDGMVGVDSSTVDLQAQ 46
 DB 1 QRTATIRGDAPAGGAVPGGAVPGGFGFGFPGVLDGMVGVDSSTVELQAQ 56
 RESULT 11
 AAE16189
 ID AAE16189 standard; peptide: 56 AA.
 AC AAE16189;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE B. bronchiseptica strain prn1 pertactin outer membrane protein region I.
 XX
 KW Pertactin; PRN; outer membrane protein; vaccine; Bordetella infection;
 KW therapy; antibiotic; antibacterial; region I.
 XX
 OS Bordetella bronchiseptica.
 XX
 PN WO200190143-A2.
 XX
 PD 29-NOV-2001.
 XX
 PF 23-MAY-2001; 2001WO-EP06457.
 XX
 PR 25-MAY-2000; 2000US-206969P.
 XX
 PA (INSP) INST PASTEUR.

XX
 PI Guiso-maclouf N, Boursaux-eude C;
 XX
 DR WPI; 2002-097639/13.
 XX
 PT Polypeptides containing polymorphisms of the repeated regions of
 PT pertactin in Bordetella species, useful in immunogenic compositions for
 PT treating infections caused by Bordetella and in diagnostic methods -
 XX
 PS Disclosure; Fig 1b; 47pp; English.
 CC The present invention relates to Bordetella bronchiseptica pertactin
 CC (outer membrane protein) or their fragments. Pertactin (PRN) is used as
 CC vaccine. Pertactin antibody is useful for treating Bordetella infections
 CC and used to detect Bordetella antigens in biological preparations or in
 CC purifying corresponding proteins, glycoproteins or their mixtures when
 CC used in affinity chromatographic columns. Pertactin is useful as antigens
 CC to identify antibodies to Bordetella in materials such as human or other
 CC animal tissue and human or other animal cells, as well as biological
 CC fluids, such as human or other animal body fluids, including human sera,
 CC and to determine the concentration of Ab in those materials. Thus the
 CC antigens can be used for qualitative or quantitative determination of
 CC Bordetella in a material. The present sequence is B. bronchiseptica
 CC pertactin outer membrane protein region I.
 CC
 SQ Sequence 56 AA;
 Query Match 86.5%; Score 212; DB 23; Length 56;
 Best Local Similarity 75.0%; Pred. No. 3,1e-18;
 Matches 42; Conservative 3; Mismatches 1; Indels 10; Gaps 1;
 QY 1 ORATTIRGDAPAGGAVPGGAV-----PGGFGPLDGMVGVDSSTVDLQAQ 46
 DB 1 QRTATIRGDAPAGGAVPGGAVPGGFGFGFPGVLDGMVGVDSSTVELQAQ 56
 RESULT 12
 AAE16191
 ID AAE16191 standard; peptide: 56 AA.
 AC AAE16191;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE B. bronchiseptica strain prn3 pertactin outer membrane protein region I.
 XX
 KW Pertactin; PRN; outer membrane protein; vaccine; Bordetella infection;
 KW therapy; antibiotic; antibacterial; region I.
 XX
 OS Bordetella bronchiseptica.
 XX
 PN WO200190143-A2.
 XX
 PD 29-NOV-2001.
 XX
 PF 23-MAY-2001; 2001WO-EP06457.
 XX
 PR 25-MAY-2000; 2000US-206969P.
 XX
 PA (INSP) INST PASTEUR.
 PI Guiso-maclouf N, Boursaux-eude C;
 XX
 DR WPI; 2002-097639/13.
 XX
 PT Polypeptides containing polymorphisms of the repeated regions of
 PT pertactin in Bordetella species, useful in immunogenic compositions for
 PT treating infections caused by Bordetella and in diagnostic methods -
 XX
 PS Disclosure; Fig 1b; 47pp; English.
 CC The present invention relates to Bordetella bronchiseptica pertactin
 CC (outer membrane protein) or their fragments. Pertactin (PRN) is used as

OY 1 QRATIRRGDAPAGAVPGAV-----PGGFGPLLDGMYGVDSSTVDLAQ 46
 DB 254 QRATIRRGDAPAGAVPGAVPGAVPGGFGFGFVLDGMYGVDSSTVELAQ 309

RESULT 15

AAE16190
 ID AAE16190 standard; peptide; 61 AA.

AC AAE16190;

DT 26-MAR-2002 (first entry)

DE B. bronchiseptica strain prn2 pertactin outer membrane protein region I.

KW Pertactin; PRN; outer membrane protein; vaccine; Bordetella infection;

OS Bordetella bronchiseptica.

PN W0200190143-A2.

PD 29-NOV-2001.

PF 23-MAY-2001; 2001WO-EP06457.

PR 25-MAY-2000; 2000US-206969P.

PA (INSP) INST PASTEUR.

PI Guiso-maclouf N, Boursaux-eude C;

DR WPI; 2002-097639/13.

PT Polypeptides containing polymorphisms of the repeated regions of
 PT pertactin in Bordetella species, useful in immunogenic compositions for
 PT treating infections caused by Bordetella and in diagnostic methods

PS Disclosure; Fig 1b; 47pp; English.

CC The present invention relates to Bordetella bronchiseptica pertactin
 CC (outer membrane protein) or their fragments. Pertactin (PRN) is used as
 CC vaccine. Pertactin antibody is useful for treating Bordetella infections
 CC and used to detect Bordetella antigens in biological preparations or in
 CC purifying corresponding proteins, glycoproteins or their mixtures when
 CC used in affinity chromatographic columns. Pertactin is useful as antigens
 CC to identify antibodies to Bordetella in materials such as human or other
 CC animal tissue and human or other animal body fluids, including human sera,
 CC fluids, such as human or other animal body fluids, including human sera,
 CC and to determine the concentration of Ab in those materials. Thus the
 CC antigens can be used for qualitative or quantitative determination of
 CC Bordetella in a material. The present sequence is B. bronchiseptica
 CC pertactin outer membrane protein region I.

SQ Sequence 61 AA;

Query Match

Best Local Similarity 85.5%; Score 209.5; DB 23; Length 61;

Matches 42; Conservative 3; Mismatches 1; Indels 15; Gaps 1;

OY 1 QRATIRRGDAPAGAVPGAV-----PGGFGPLLDGMYGVDSSTVDLAQ 45
 DB 1 QRATIRRGDAPAGAVPGAVPGGFGFGFVLDGMYGVDSSTVELA 60
 OY 46 Q 46
 DB 61 Q 61

Search completed: March 17, 2003, 11:37:07
 Job time : 61.4314 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 17, 2003, 11:38:18 ; Search time 19.8431 Seconds
(Without alignments)
106.850 Million cell updates/sec

Title: US-09-855-754b-8

Perfect score: 245
Sequence: 1 ORATIRRGDAPAGAVPGGA.....PLDGMVGVDSSTVDLAQ 46

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 199416 seqs, 46092074 residues

Total number of hits satisfying chosen parameters: 199416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA:*

1: /cgn2_6/ptodata/1/pubppaa/US08_NEWL_PUB.pep:*
2: /cgn2_6/ptodata/1/pubppaa/PCITUS_NEWL_PUB.pep:*
3: /cgn2_6/ptodata/1/pubppaa/US06_NEWL_PUB.pep:*
4: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubppaa/US07_NEWL_PUB.pep:*
6: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubppaa/PCITUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubppaa/US09_NEWL_PUB.pep:*
10: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubppaa/US10_NEWL_PUB.pep:*
12: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubppaa/US60_NEWL_PUB.pep:*
14: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	245	100.0	46	9	US-09-855-754-8
2	245	100.0	911	9	US-08-855-754-4
3	232.5	94.9	51	9	US-09-855-754-7
4	232.5	94.9	922	9	US-09-855-754-6
5	219	89.4	910	9	US-09-855-754-5
6	214.5	87.6	51	9	US-09-855-754-13
7	212	86.5	56	9	US-09-855-754-9
8	212	86.5	56	9	US-09-855-754-10
9	212	86.5	56	9	US-09-855-754-12
10	209.5	85.5	61	9	US-09-855-754-11
11	60	24.5	450	10	US-09-812-382-6
12	60	24.5	457	9	US-10-078-770-194
13	59	24.1	50	10	US-09-812-382-3
14	58	23.7	178	10	US-09-815-242-10448
15	57	23.3	564	9	US-10-099-619-2
16	56.5	23.1	348	10	US-09-918-568-50
17	56.5	23.1	574	10	US-09-782-980-15
18	56.5	23.1	574	10	US-09-909-743-6
19	56	22.9	119	10	US-09-861-597-9

20	56	22.9	284	9	US-10-063-547-62	Sequence 62, Appl
21	56	22.9	284	9	US-10-174-580-236	Sequence 236, App
22	56	22.9	284	9	US-10-176-758-236	Sequence 236, App
23	56	22.9	284	9	US-10-063-616-62	Sequence 62, Appl
24	56	22.9	284	9	US-10-175-737-236	Sequence 236, App
25	56	22.9	284	9	US-10-063-502-62	Sequence 62, Appl
26	56	22.9	284	9	US-10-173-706-236	Sequence 236, App
27	56	22.9	284	9	US-10-175-738-236	Sequence 236, App
28	56	22.9	284	9	US-10-175-732-236	Sequence 236, App
29	56	22.9	284	9	US-10-176-482-236	Sequence 236, App
30	56	22.9	284	9	US-10-176-757-236	Sequence 236, App
31	56	22.9	284	9	US-10-176-913-236	Sequence 236, App
32	56	22.9	284	9	US-10-180-552-236	Sequence 236, App
33	56	22.9	284	9	US-10-180-557-236	Sequence 236, App
34	56	22.9	284	9	US-10-173-700-236	Sequence 236, App
35	56	22.9	284	9	US-10-174-572-236	Sequence 236, App
36	56	22.9	284	9	US-10-174-579-236	Sequence 236, App
37	56	22.9	284	9	US-10-174-582-236	Sequence 236, App
38	56	22.9	284	9	US-10-174-588-236	Sequence 236, App
39	56	22.9	284	9	US-10-175-739-236	Sequence 236, App
40	56	22.9	284	9	US-10-175-740-236	Sequence 236, App
41	56	22.9	284	9	US-10-175-743-236	Sequence 236, App
42	56	22.9	284	9	US-10-176-488-236	Sequence 236, App
43	56	22.9	284	9	US-10-176-492-236	Sequence 236, App
44	56	22.9	284	9	US-10-176-747-236	Sequence 236, App
45	56	22.9	284	9	US-10-176-750-236	Sequence 236, App

ALIGNMENTS

RESULT 1
US-09-855-754-8
Sequence 8, Application US/09855754
Publication No. US20020192237A1
GENERAL INFORMATION:
APPLICANT: BOUSSAUX-ETUDE, CAROLINE
APPLICANT: GUISO-MACLOUF, NICOLE
TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED
TITLE OF INVENTION: REGIONS OF PERTACTIN IN BORDETELLA PERTUSSIS.
TITLE OF INVENTION: BRONCHISEPTICA, THEIR USE IN DIAGNOSTICS, AND IN
TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS
FILE REFERENCE: 03495-0206-00000
CURRENT APPLICATION NUMBER: US/09/855,754
PRIOR APPLICATION NUMBER: 2001-09-10
PRIOR FILING DATE: 2000-05-25
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 8
LENGTH: 46
TYPE: PRT
ORGANISM: Bordetella bronchiseptica
US-09-855-754-8
Query Match 100.0%; Score 245; DB 9; Length 46;
Best Local Similarity 100.0%; Pred. No. 2.8e-21;
Matches 46; Conservative 0; Mismatches 0; Indels 0;
Gaps 0;
Gy 1 ORATIRRGDAPAGAVPGGPGPLDGMVGVDSSTVDLAQ 46
Db 1 ORATIRRGDAPAGAVPGGPGPLDGMVGVDSSTVDLAQ 46
RESULT 2
US-09-855-754-4
Sequence 4, Application US/09855754
Publication No. US20020192237A1
GENERAL INFORMATION:
APPLICANT: BOUSSAUX-ETUDE, CAROLINE
APPLICANT: GUISO-MACLOUF, NICOLE
TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED

;; TITLE OF INVENTION: REGIONS OF PERTACTIN IN BORDETELLA PERTUSSIS,
;; TITLE OF INVENTION: BORDETELLA PARAPERTUSSIS, AND BORDETELLA
;; TITLE OF INVENTION: BRONCHISEPTICA, THEIR USE IN DIAGNOSTICS, AND IN
;; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS
;; FILE REFERENCE: 03495-0206-00000
;; CURRENT APPLICATION NUMBER: US/09/855,754
;; PRIOR FILING DATE: 2001-09-10
;; PRIOR APPLICATION NUMBER: 60/206,969
;; PRIOR FILING DATE: 2000-05-25
;; NUMBER OF SEQ ID NOS: 24
;; SOFTWARE: Patentln Ver. 2.1
;; SEQ ID NO 4
;; LENGTH: 911
;; TYPE: PRF
;; ORGANISM: Bordetella bronchiseptica
US-09-855-754-4

Query Match 100.0%; Score 245; DB 9; Length 911;
Best Local Similarity 100.0%; Pred. No. 5.8e-20;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ORATIRRGDAPAGGAVPGGFGPPLDGMVGVDSSTVDLAQ 46
Db 254 ORATIRRGDAPAGGAVPGGAVPGGFGPPLDGMVGVDSSTVDLAQ 299

RESULT 3
US-09-855-754-7
;; Sequence 7, Application US/09855754
;; Publication No. US20020192237A1
;; GENERAL INFORMATION:
;; APPLICANT: BOURSAX-EDDE, CAROLINE
;; APPLICANT: GUISSO-MACLOUF, NICOLE
;; TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED
;; TITLE OF INVENTION: REGIONS OF PERTACTIN IN BORDETELLA PERTUSSIS,
;; TITLE OF INVENTION: BORDETELLA PARAPERTUSSIS, AND BORDETELLA
;; TITLE OF INVENTION: BRONCHISEPTICA, THEIR USE IN DIAGNOSTICS, AND IN
;; FILE REFERENCE: 03495-0206-00000
;; CURRENT APPLICATION NUMBER: US/09/855,754
;; CURRENT FILING DATE: 2001-09-10
;; PRIOR APPLICATION NUMBER: 60/206,969
;; PRIOR FILING DATE: 2000-05-25
;; NUMBER OF SEQ ID NOS: 24
;; SOFTWARE: Patentln Ver. 2.1
;; SEQ ID NO 7
;; LENGTH: 51
;; TYPE: PRF
;; ORGANISM: Bordetella bronchiseptica
US-09-855-754-7

Query Match 94.9%; Score 232.5; DB 9; Length 51;
Best Local Similarity 90.2%; Pred. No. 7.9e-20;
Matches 46; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

OY 1 ORATIRRGDAPAGGAVPGGFGPPLDGMVGVDSSTVDLAQ 46
Db 1 ORATIRRGDAPAGGAVPGGAVPGGFGPPLDGMVGVDSSTVDLAQ 51

RESULT 4
US-09-855-754-6
;; Sequence 6, Application US/09855754
;; Publication No. US20020192237A1
;; GENERAL INFORMATION:
;; APPLICANT: BOURSAX-EDDE, CAROLINE
;; APPLICANT: GUISSO-MACLOUF, NICOLE
;; TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED
;; TITLE OF INVENTION: REGIONS OF PERTACTIN IN BORDETELLA PERTUSSIS,
;; TITLE OF INVENTION: BORDETELLA PARAPERTUSSIS, AND BORDETELLA
;; TITLE OF INVENTION: BRONCHISEPTICA, THEIR USE IN DIAGNOSTICS, AND IN
;; FILE REFERENCE: 03495-0206-00000

;; CURRENT APPLICATION NUMBER: US/09/855,754
;; CURRENT FILING DATE: 2001-09-10
;; PRIOR APPLICATION NUMBER: 60/206,969
;; PRIOR FILING DATE: 2000-05-25
;; NUMBER OF SEQ ID NOS: 24
;; SOFTWARE: Patentln Ver. 2.1
;; SEQ ID NO 6
;; LENGTH: 922
;; TYPE: PRF
;; ORGANISM: Bordetella parapertussis
US-09-855-754-6

Query Match 94.9%; Score 232.5; DB 9; Length 922;
Best Local Similarity 90.2%; Pred. No. 1.5e-18;
Matches 46; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

OY 1 ORATIRRGDAPAGGAVPGGFGPPLDGMVGVDSSTVDLAQ 46
Db 254 ORATIRRGDAPAGGAVPGGAVPGGFGPPLDGMVGVDSSTVDLAQ 304

RESULT 5
US-09-855-754-5
;; Sequence 5, Application US/09855754
;; Publication No. US20020192237A1
;; GENERAL INFORMATION:
;; APPLICANT: BOURSAX-EDDE, CAROLINE
;; APPLICANT: GUISSO-MACLOUF, NICOLE
;; TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED
;; TITLE OF INVENTION: REGIONS OF PERTACTIN IN BORDETELLA PERTUSSIS,
;; TITLE OF INVENTION: BORDETELLA PARAPERTUSSIS, AND BORDETELLA
;; TITLE OF INVENTION: BRONCHISEPTICA, THEIR USE IN DIAGNOSTICS, AND IN
;; FILE REFERENCE: 03495-0206-00000
;; CURRENT APPLICATION NUMBER: US/09/855,754
;; CURRENT FILING DATE: 2001-09-10
;; PRIOR APPLICATION NUMBER: 60/206,969
;; PRIOR FILING DATE: 2000-05-25
;; NUMBER OF SEQ ID NOS: 24
;; SOFTWARE: Patentln Ver. 2.1
;; SEQ ID NO 5
;; LENGTH: 910
;; TYPE: PRF
;; ORGANISM: Bordetella pertussis
US-09-855-754-5

Query Match 89.4%; Score 219; DB 9; Length 910;
Best Local Similarity 76.8%; Pred. No. 4.8e-17;
Matches 43; Conservative 3; Mismatches 0; Indels 10; Gaps 1;

OY 1 ORATIRRGDAPAGGAVPGGAVPGGFGPPLDGMVGVDSSTVDLAQ 46
Db 254 ORATIRRGDAPAGGAVPGGAVPGGFGPPLDGMVGVDSSTVDLAQ 309

RESULT 6
US-09-855-754-13
;; Sequence 13, Application US/09855754
;; Publication No. US20020192237A1
;; GENERAL INFORMATION:
;; APPLICANT: BOURSAX-EDDE, CAROLINE
;; APPLICANT: GUISSO-MACLOUF, NICOLE
;; TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED
;; TITLE OF INVENTION: REGIONS OF PERTACTIN IN BORDETELLA PERTUSSIS,
;; TITLE OF INVENTION: BORDETELLA PARAPERTUSSIS, AND BORDETELLA
;; TITLE OF INVENTION: BRONCHISEPTICA, THEIR USE IN DIAGNOSTICS, AND IN
;; FILE REFERENCE: 03495-0206-00000
;; CURRENT APPLICATION NUMBER: US/09/855,754
;; CURRENT FILING DATE: 2001-09-10
;; PRIOR APPLICATION NUMBER: 60/206,969
;; PRIOR FILING DATE: 2000-05-25
;; NUMBER OF SEQ ID NOS: 24

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 13

LENGTH: 51

TYPE: PRT

ORGANISM: Bordetella bronchiseptica

US-09-855-754-13

Query Match 87.6%; Score 214.5; DB 9; Length 51;

Best Local Similarity 82.4%; Pred. No. 8.3e-18;

Matches 42; Conservative 3; Mismatches 1; Indels 5; Gaps 1;

1 QARTRRGDAPAGAVPGAVPGGFGPVLDMGWYDVSDSTVDLAQ 46

1 QARTRRGDAPAGAVPGAVPGGFGPVLDMGWYDVSDSTVDLAQ 51

US-09-855-754-9

Sequence 9, Application US/09855754

Publication No. US20020192237A1

GENERAL INFORMATION:

APPLICANT: BOURSAX-EUDE, CAROLINE

APPLICANT: GUISSO-MACLOUF, NICOLE

TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED

TITLE OF INVENTION: REGIONS OF PERTACTIN IN BORDETELLA PERTUSSIS,

TITLE OF INVENTION: BORDETELLA PARAPERTUSSIS, AND BORDETELLA

TITLE OF INVENTION: BRONCHISEPTICA, THEIR USE IN DIAGNOSTICS, AND IN

FILE REFERENCE: 03495-0206-00000

CURRENT APPLICATION NUMBER: US/09/855,754

CURRENT FILING DATE: 2001-09-10

PRIOR APPLICATION NUMBER: 60/206,969

PRIOR FILING DATE: 2000-05-25

NUMBER OF SEQ ID NOS: 24

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 9

LENGTH: 56

TYPE: PRT

ORGANISM: Bordetella bronchiseptica

US-09-855-754-9

Query Match 86.5%; Score 212; DB 9; Length 56;

Best Local Similarity 75.0%; Pred. No. 1.7e-17;

Matches 42; Conservative 2; Mismatches 2; Indels 10; Gaps 1;

1 QARTRRGDAPAGAVPGAVPGGFGPVLDMGWYDVSDSTVDLAQ 46

1 QARTRRGDAPAGAVPGAVPGGFGPVLDMGWYDVSDSTVDLAQ 56

US-09-855-754-10

Sequence 10, Application US/09855754

Publication No. US20020192237A1

GENERAL INFORMATION:

APPLICANT: BOURSAX-EUDE, CAROLINE

APPLICANT: GUISSO-MACLOUF, NICOLE

TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED

TITLE OF INVENTION: REGIONS OF PERTACTIN IN BORDETELLA PERTUSSIS,

TITLE OF INVENTION: BORDETELLA PARAPERTUSSIS, AND BORDETELLA

TITLE OF INVENTION: BRONCHISEPTICA, THEIR USE IN DIAGNOSTICS, AND IN

FILE REFERENCE: 03495-0206-00000

CURRENT APPLICATION NUMBER: US/09/855,754

CURRENT FILING DATE: 2001-09-10

PRIOR APPLICATION NUMBER: 60/206,969

PRIOR FILING DATE: 2000-05-25

NUMBER OF SEQ ID NOS: 24

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 10

LENGTH: 56

TYPE: PRT

ORGANISM: Bordetella bronchiseptica

US-09-855-754-10

Query Match 86.5%; Score 212; DB 9; Length 56;

Best Local Similarity 75.0%; Pred. No. 1.7e-17;

Matches 42; Conservative 3; Mismatches 1; Indels 10; Gaps 1;

1 QARTRRGDAPAGAVPGAVPGGFGPVLDMGWYDVSDSTVDLAQ 46

1 QARTRRGDAPAGAVPGAVPGGFGPVLDMGWYDVSDSTVDLAQ 56

US-09-855-754-12

Sequence 12, Application US/09855754

Publication No. US20020192237A1

GENERAL INFORMATION:

APPLICANT: BOURSAX-EUDE, CAROLINE

APPLICANT: GUISSO-MACLOUF, NICOLE

TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED

TITLE OF INVENTION: REGIONS OF PERTACTIN IN BORDETELLA PERTUSSIS,

TITLE OF INVENTION: BORDETELLA PARAPERTUSSIS, AND BORDETELLA

TITLE OF INVENTION: BRONCHISEPTICA, THEIR USE IN DIAGNOSTICS, AND IN

FILE REFERENCE: 03495-0206-00000

CURRENT APPLICATION NUMBER: US/09/855,754

CURRENT FILING DATE: 2001-09-10

PRIOR APPLICATION NUMBER: 60/206,969

PRIOR FILING DATE: 2000-05-25

NUMBER OF SEQ ID NOS: 24

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 12

LENGTH: 56

TYPE: PRT

ORGANISM: Bordetella bronchiseptica

US-09-855-754-12

Query Match 86.5%; Score 212; DB 9; Length 56;

Best Local Similarity 75.0%; Pred. No. 1.7e-17;

Matches 42; Conservative 3; Mismatches 1; Indels 10; Gaps 1;

1 QARTRRGDAPAGAVPGAVPGGFGPVLDMGWYDVSDSTVDLAQ 46

1 QARTRRGDAPAGAVPGAVPGGFGPVLDMGWYDVSDSTVDLAQ 56

US-09-855-754-11

Sequence 11, Application US/09855754

Publication No. US20020192237A1

GENERAL INFORMATION:

APPLICANT: BOURSAX-EUDE, CAROLINE

APPLICANT: GUISSO-MACLOUF, NICOLE

TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED

TITLE OF INVENTION: REGIONS OF PERTACTIN IN BORDETELLA PERTUSSIS,

TITLE OF INVENTION: BORDETELLA PARAPERTUSSIS, AND BORDETELLA

TITLE OF INVENTION: BRONCHISEPTICA, THEIR USE IN DIAGNOSTICS, AND IN

FILE REFERENCE: 03495-0206-00000

CURRENT APPLICATION NUMBER: US/09/855,754

CURRENT FILING DATE: 2001-09-10

PRIOR APPLICATION NUMBER: 60/206,969

PRIOR FILING DATE: 2000-05-25

NUMBER OF SEQ ID NOS: 24

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 11

LENGTH: 61

TYPE: PRT

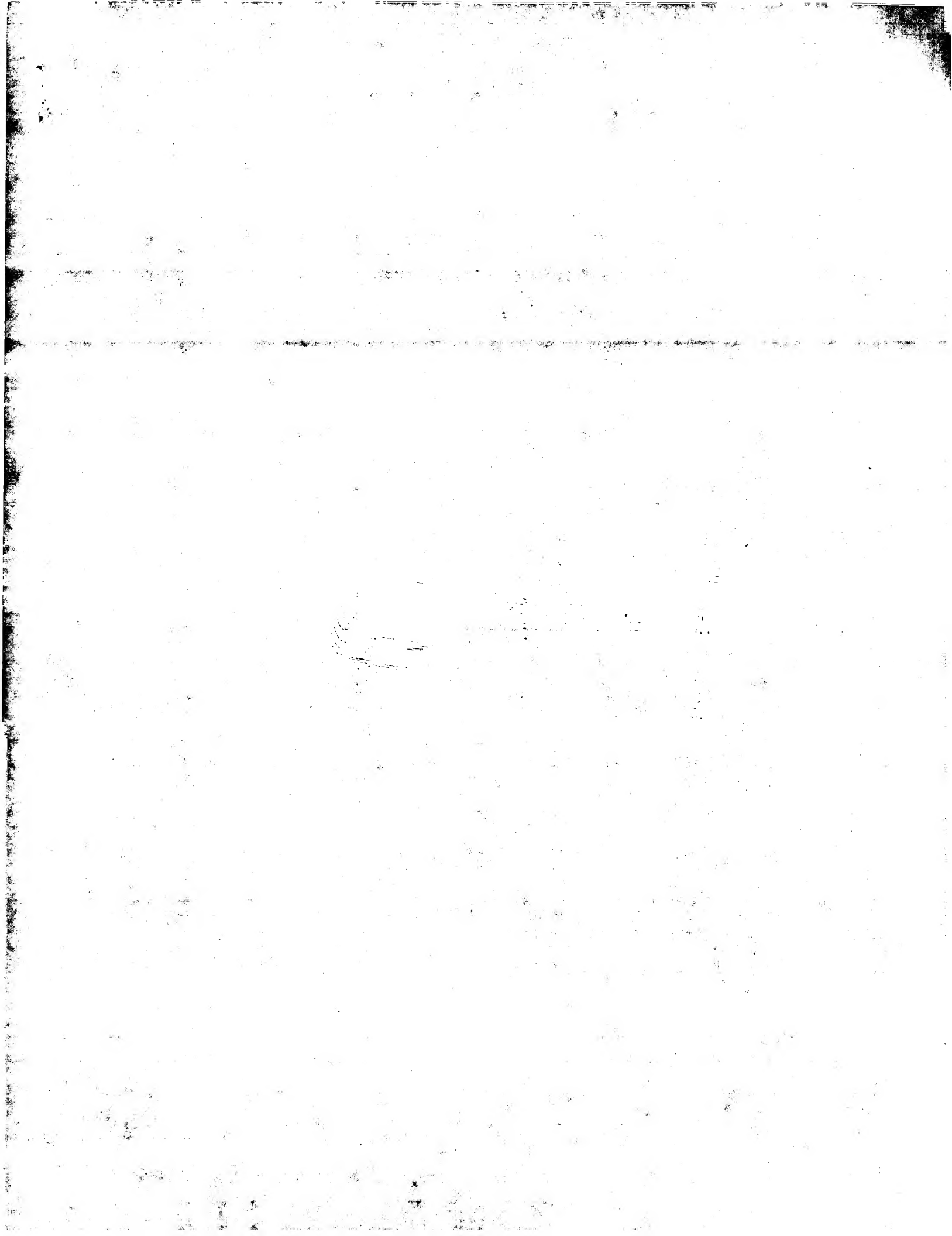
ORGANISM: Bordetella bronchiseptica

US-09-855-754-11

Query Match 85.5%; Score 209.5; DB 9; Length 61;

Best Local Similarity 68.9%; Pred. No. 3.6e-17;

Matches 42; Conservative 3; Mismatches 1; Indels 15; Gaps 1;



GenCore version 5.1.4-EP5-4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 17, 2003, 11:35:03 ; Search time 25.2549 Seconds
(without alignments)
175.102 Million cell updates/sec

Title: US-09-855-754B-8

Sequence: 1 GRATIRRGDAPAGAVPGCA.....PLDDGWTGVSDSTVDLAQ 46

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: PIR-73:*
2: PIR1:*
3: PIR2:*
4: PIR3:*
5: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Query Length	DB ID	Description
1	245	100.0	911	2 A47675	68K outer membrane
2	232.5	94.9	922	2 S15204	pertactin - Bordet
3	202	82.4	910	2 A32550	outer membrane pro
4	73.5	30.0	646	2 T43677	ATP-dependent RNA
5	68	27.8	641	2 PC7036	heat shock protein
6	67	27.3	783	2 T35389	probable serine-th
7	64.5	26.3	565	1 HMIYVN	hemagglutinin prec
8	64	26.1	645	2 S41372	dnak-type molecula
9	62	25.3	524	2 AG3328	proteinase do (EC
10	61.5	25.1	514	2 F87552	hypothetical prote
11	61.5	25.1	1300	2 T03186	probable immediate
12	61	24.9	114	2 C98152	NAD-dependent form
13	61	24.9	114	2 A13135	hypothetical prote
14	60	24.5	215	2 T22572	hypothetical prote
15	60	24.5	231	2 S37108	hypothetical prote
16	60	24.5	424	2 T29127	cuticlin 2 - Caeno
17	59.5	24.3	566	2 S69867	hypothetical prote
18	59.5	24.3	860	1 EAMS	hemagglutinin prec
19	59	24.1	185	2 T49890	elastin precursor
20	59	24.1	294	2 S13141	glycine-rich prote
21	59	24.1	333	2 AB3237	hypothetical prote
22	59	24.1	405	2 T29167	hypothetical prote
23	59	24.1	420	2 T59234	octamer binding tr
24	59	24.1	472	2 B56954	yes-associated pro
25	59	24.1	513	2 D70836	probable PPE prote
26	59	24.1	566	1 HMIYD8	hemagglutinin prec
27	59	24.1	1711	1 A55148	protein-tyrosine-p
28	58.5	23.9	128	2 S25103	MSIE protein - ma
29	58.5	23.9	480	1 B70804	DNA repair protein

30	58.5	23.9	5627	2 C83339	hypothetical prote
31	58	23.7	178	1 DDEC	single-stranded DN
32	58	23.7	178	2 A91259	ssDNA-binding prot
33	58	23.7	178	2 E86099	ssDNA-binding prot
34	58	23.7	244	2 T13580	hypothetical prote
35	58	23.7	304	2 H83636	hypothetical prote
36	58	23.7	313	2 A15236	conserved hypotet
37	58	23.7	526	2 AF3004	conserved hypotet
38	58	23.7	526	2 C98279	hypothetical prote
39	58	23.7	577	2 A40220	cleavage stimulat
40	58	23.7	784	2 A26601	elastin precursor
41	57.5	23.5	529	2 J00783	55.5K sporulation
42	57.5	23.5	565	2 HMIY	hemagglutinin prec
43	57.5	23.5	566	1 HMIY	hemagglutinin prec
44	57.5	23.5	566	1 HMIYVR	hemagglutinin prec
45	57.5	23.5	566	1 HMIYVN	hemagglutinin prec

ALIGNMENTS

RESULT 1

A47675
68K outer membrane protein P.68 pertactin - Bordetella bronchiseptica
C:Species: Bordetella bronchiseptica
C>Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 08-Oct-1999
C:Accession: A47675
R:LI, J.; Fairweather, N.F.; Novotny, P.; Dougan, G.; Charles, I.G.
J. Gen. Microbiol. 138, 1697-1705, 1992
A:Title: Cloning, nucleotide sequence and heterologous expression of the protective o
A:Reference number: A47675; MUID:92407514; PMID:1527510
A:Contents: CN7531
A:Accession: A47675
A:Status: Preliminary
A:Molecule type: nucleic acid
A:Residues: 1-911
A:Cross-references: GB:X54815; GB:S46416; NID:939396; PIDN:CA38584.1; PID:939397
A>Note: sequence extracted from NCBI backbone (NCBI:113318, NCBI:P.113319)

Query Match

Best Local Similarity 100.0%; Score 245; DB 2; Length 911;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GRATIRRGDAPAGAVPGCAVPGGFPPLDDGWTGVSDSTVDLAQ 46

DB 254 GRATIRRGDAPAGAVPGCAVPGGFPPLDDGWTGVSDSTVDLAQ 299

RESULT 2

S15204
pertactin - Bordetella parapertussis

N:Alternate names: outer membrane protein P70
C:Species: Bordetella parapertussis

C>Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 08-Oct-1999
C:Accession: S15204; S14659

R:LI, L.J.; Dougan, G.; Novotny, P.; Charles, I.G.
Mol. Microbiol. 5, 409-417, 1991

A:Title: P.70 pertactin an outer-membrane protein from Bordetella parapertussis: c10
A:Reference number: S15204; MUID:91251771; PMID:2041476

A:Accession: S15204

A:Molecule type: DNA

A:Residues: 1-922

A:Cross-references: EMBL:X54547; NID:939761; PIDN:CA38419.1; PID:939762

A:Gene: prn

C:Keywords: membrane protein

Query Match 94.9%; Score 232.5; DB 2; Length 922;
Best Local Similarity 90.2%; Pred. No. 1.1e-18;
Matches 46; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

OY 1 GRATIRRGDAPAGAVPGCAVPGGFPPLDDGWTGVSDSTVDLAQ 46

Db 254 ORATIRRGDAPAGAVPGGAVPGGAVPGGFGPLLDGMVGVDSSTVDLAQ 304

RESULT 3

A32560

outer membrane protein P.69 precursor - Bordetella pertussis

C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 08-Oct-1999

C:Accession: A32560

R:Charles, I.G.; Dougan, G.; Pickard, D.; Chatfield, S.; Smith, M.; Novotny, P.; Morris

Proc. Natl. Acad. Sci. U.S.A. 86, 3554-3558, 1989

A:Title: Molecular cloning and characterization of protective outer membrane protein P.69

A:Reference number: A32560, MUID:69264462, PMID:2542937

A:Accession: A32560

A:Molecule type: DNA

A:Residues: 1-910 <CH>

A:Cross-references: GB:J04560; NID:q144053; PIDN:AAA22980.1; PID:q144054

A:Note: It is uncertain whether Met-1 or Met-3 is the initiator

C:Keywords: membrane protein

F:35-910/Product: outer membrane protein P.69 status predicted <MAT>

Query Match 82.4%; Score 202; DB 2; Length 910;

Best Local Similarity 73.2%; Pred. No. 3, 3e-15;

Matches 41; Conservative 3; Mismatches 2; Indels 10; Gaps 1;

Db 254 ORATIRRGDAPAGAVPGGAVPGGAVPGGFGPLLDGMVGVDSSTVDLAQ 309

RESULT 4

T45677

ATP-dependent RNA helicase-like protein - Arabidopsis thaliana

N:Alternate names: protein Flap22.160

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 17-Nov-2000

C:Accession: T45677

R:D'Angelio, M.; Vezzi, A.; Modesto, D.; Pigazzi, M.; Valle, G.; Mewes, H.W.; Lemcke, K.

submitted to the Protein Sequence Database, January 2000

A:Reference number: Z23011

A:Accession: T45677

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-646 <DN>

A:Cross-references: EMBL:AL137082

A:Experimental source: cultivar Columbia; BAC clone Flap22

C:Genetics:

A:Map position: 3

A:Introns: 239/3; 267/3; 348/3; 404/3; 442/3

A:Note: Flap22.160

C:Superfamily: ATP-dependent RNA helicase DBP1

Query Match 30.0%; Score 73.5; DB 2; Length 646;

Best Local Similarity 40.0%; Pred. No. 1;

Matches 18; Conservative 5; Mismatches 11; Indels 11; Gaps 2;

Db 562 RRSFPRGCGGADYGGGGGCGGCGTGMFGGYPVGGG 606

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-641 <CER>

A:Cross-references: GB:AF188289

C:Superfamily: heat shock protein 70

Query Match 27.8%; Score 68; DB 2; Length 641;

Best Local Similarity 61.1%; Pred. No. 4.3;

Matches 11; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Db 609 GGMGCGGMPGCGAGPGE 626

RESULT 6

T33589

probable serine-threonine protein kinase - Streptomyces coelicolor

C:Species: Streptomyces coelicolor

C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 21-Jan-2000

C:Accession: T33589

R:Murphy, L.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, June 1999

A:Reference number: Z21576

A:Accession: T33589

A>Status: preliminary; translated from GB/EMBL/DBD

A:Molecule type: DNA

A:Residues: 1-783 <MR>

A:Cross-references: EMBL:AL079348; PIDN:CAB45488.1; GSPDB:GN00070; SCOEDB:SC66T3.32C

A:Experimental source: strain A3(2)

C:Genetics:

A:Gene: SCOEDB:SC66T3.32C

C:Superfamily: unassigned collagens

Query Match 27.3%; Score 67; DB 2; Length 783;

Best Local Similarity 61.9%; Pred. No. 6.8;

Matches 13; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Db 376 RGGVGRGAGPGRGPGG 396

RESULT 7

HMTVNT

hemagglutinin precursor - Influenza A virus (strain A/Turkey/Ontario/6118/68)

N:Contains: hemagglutinin HA1; hemagglutinin HA2

C:Species: Influenza A virus

C:Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 16-Jun-2000

C:Accession: F39987; A04067; C33157

R:Nobusawa, E.; Aoyama, T.; Kato, H.; Suzuki, Y.; Tatenio, Y.; Nakajima, K.

Virolgy 182, 475-485, 1991

A:Title: Comparison of complete amino acid sequences and receptor-binding properties

A:Reference number: A39987; MUID:91220697; PMID:2024485

A:Accession: F39987

A:Molecule type: genomic RNA

A:Residues: 1-565 <ND>

A:Cross-references: GB:D0304; NID:9221317; PIDN:BA14334.1; PID:9221318

A:Experimental source: strain A/Turkey/Ontario/6118/68 [H8N4]

A:Alt. G.M.

Proc. Natl. Acad. Sci. U.S.A. 78, 7639-7643, 1981

A:Title: Sequence relationships among the hemagglutinin genes of 12 subtypes of Infl

A:Reference number: A93902; MUID:82150925; PMID:6174976

A:Accession: A04067

A:Molecule type: genomic RNA

A:Residues: 1-8, 'W', 'L', '11-35', 'P', '37-88', 'VN', '91-99 <AIR>

A:Cross-references: GB:J02089; NID:9324129; PIDN:AAA43177.1; PID:9324130

A:Experimental source: strain A/Turkey/Ontario/6118/68 [H8]

A:Note: The signal sequence and the amino end of the HA1 chain comprise residues 1-16

C:Genetics:

A:Map position: segment 4

C:Superfamily: Influenza virus hemagglutinin

C:Keywords: glycoprotein; hemagglutinin; homotrimer; transmembrane protein

F:1-16/Domain: signal sequence #status predicted <SIG>
F:17-343/Product: hemagglutinin HA1 #status predicted <HA1>
F:344-555/Product: hemagglutinin HA2 #status predicted <HA2>
F:536-552/Domain: transmembrane #status predicted <TM>
F:26/27/139,150,222,303,310,497,524/Binding site: carbohydrate (asn) (covalent) #status

Query Match 26.1%; Score 64.5; DB 1; Length 565;
Best Local Similarity 43.8%; Pred. No. 9.4;
Matches 14; Conservative 6; Mismatches 9; Indels 3; Gaps 2;

OY 11 PAG--GAVPGAVPGGPRLLDGYGVSDS 40
DB 342 PGLFGLAI-AGFIEGWSGMIDGYGFHNSNS 372

RESULT 8

dnak-type molecular chaperone HSA1 - yeast (Pichia angusta)
N:Alternate names: heat shock protein hsp72
C:Species: Pichia angusta

C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 20-Apr-2000
C:Accession: S41372; S71640; S71636
R:Diesel, A.A.; Roggenkamp, R.R.

Submitted to the EMBL Data Library, January 1994

A:Description: HSP70 genes of the yeast Hansenula polymorpha.

A:Reference number: S41372

A:Accession: S41372

A:Molecule type: DNA

A:Residues: 1-645 <DIE>

A:Cross-references: EMBL:Z29379; NID:9443914; PIDN:CA82570.1; PID:9443915

R:Litorenko, V.I.; Evers, M.E.; Diesel, A.; Samyn, B.; Van Beumen, J.; Roggenkamp, R.;

Yeast 12, 849-857, 1996

A:Title: Identification and characterization of cytosolic Hansenula polymorpha proteins

A:Reference number: S71636; MUID:56437974; PMID:8840502

A:Accession: S71640

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-645 <TIT>

A:Cross-references: EMBL:Z29379; NID:9443914; PIDN:CA82570.1; PID:9443915

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1994

A:Accession: S71636

A:Molecule type: protein

A:Residues: 4-10;137-143;158-164;221-227;329-335;362-368;388-394;425-431;573-579;609-615

C:Genetics:

A:Gene: HSA1; HSP72

C:Function:

A:Description: involved in protein folding and assembling/disassembling of protein comp

C:Superfamily: heat shock protein 70

C:Keywords: ATP; cytosol; heat shock; molecular chaperone; stress-induced protein

Query Match 26.1%; Score 64; DB 2; Length 645;
Best Local Similarity 70.6%; Pred. No. 12;
Matches 12; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 9 DAPGAGVPGGAVPGGF 25
DB 609 DLVAGGVPGGAPGPF 625

RESULT 9

AG3328
protease do (EC 3.4.21.-) [Imported] - Brucella melitensis (strain 16M)

C:Species: Brucella melitensis

C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002

C:Accession: AG3328

R:DelVecchio, V.G.; Kapral, V.; Redkar, R.J.; Patra, G.; Mijer, C.; Los, T.; Ivanova,

Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitens

A:Reference number: AD3252; PMID:11756688

A:Accession: AG3328

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-524 <KRR>
A:Cross-references: GB:AE008917; PIDN:AL51794.1; PID:917982337; GSPDB:GN00190

A:Experimental source: strain 16M

C:Genetics:

A:Gene: BMEI0613

A:Map position: I

C:Keywords: hydrolase; serine proteinase

Query Match 25.3%; Score 62; DB 2; Length 524;
Best Local Similarity 32.7%; Pred. No. 17;
Matches 16; Conservative 8; Mismatches 13; Indels 12; Gaps 2;

OY 10 APAGAVP--GGAVPG-----GGPRLDGYGVSDSTVDLAQ 46
DB 253 SPSSGSGIGFPAIPAMAGVYDLEKEFEVRRGWLGVLPVEDIAQ 301

RESULT 10

hypothetical protein CC2774 [Imported] - Caulobacter crescentus

C:Species: Caulobacter crescentus

C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001

C:Accession: F87592

R:Miernum, W.C.; Deboy, R.T.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg,

n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A:Title: Complete Genome Sequence of Caulobacter crescentus.

A:Reference number: A87249; MUID:21173698; PMID:11259647

A:Accession: F87592

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-514 <STO>

A:Cross-references: GB:AE005673; NID:913424372; PIDN:AK24738.1; GSPDB:GN00148

C:Genetics:

A:Gene: CC2774

Query Match 25.1%; Score 61.5; DB 2; Length 514;
Best Local Similarity 42.0%; Pred. No. 19;
Matches 21; Conservative 2; Mismatches 14; Indels 13; Gaps 3;

OY 2 RATTIRGDPAPAG--GAVPGGA--VPGGFG-----PLLDGYGVSDS 38
DB 382 RGLLRGKGPASDGAAPGAGGPGGRRLYHRLGARNPDHAGTGDLS 431

RESULT 11

T03166
probable immediate early protein - alcelaphine herpesvirus 1

C:Species: alcelaphine herpesvirus 1

C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 24-Nov-1999

C:Accession: T03166

R:Emsher, A.; Pilanz, R.; Fleckenstein, B.

J. Virol. 71, 6517-6525, 1997

A:Title: Primary structure of the alcelaphine herpesvirus 1 genome.

A:Reference number: 214840; MUID:97404659; PMID:9261371

A:Accession: T03166

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1300 <ENS>

A:Cross-references: EMBL:AF005370; NID:92337967; PIDN:AC58118.1; PID:92338034

C:Superfamily: collagen alpha 2(I) chain; fibrillar collagen carboxyl-terminal homo

Query Match 25.1%; Score 61.5; DB 2; Length 1300;
Best Local Similarity 66.7%; Pred. No. 49;
Matches 12; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

OY 8 GDAPAGAVPGGA--VPGG 24
DB 102 GEGPGGGEVPGGGEVPGG 119

RESULT 12

```

A:Residues: 1-215 <MIL>
A:Cross-references: EMBL:Z81088; PIDN:CAB03133.1; GSPDB:GN00023; CESP:F53F1.5
A:Experimental source: clone F53F1
C:Genetics:
  A:Gene: CESP:F53F1.5
  A:Map position: 5
  A:Introns: 30/2
C:Superfamily: proline-rich protein

Query Match          24.5%  Score 60;  DB 2;  length 215;
Best Local Similarity 50.0%  Pred. No. 11;
Matches 13;  Conservative 1;  Mismatches 12;  Indels 0;  Gaps 0

      0Y      8  GDAPAGCAVPGGAVPDGFGPILIDGWY  33
      |||||  |||||  |||||  |||||  |||||
Db    136  GAAPAGGAYQSGPAFGAAPAVGAY  161

```

RESULT 15
S37108
Cuticlin 2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 21-Jan-2000
C:Accession: S37108
R:Sebastiano, M.; Zel, F.; Lassandro, F.; Nola, M.; Ristrotore, F.; Bazzicalupo, P.
Submitted to the EMBL Data Library, September 1993

```

A:Reference number: S37108
A:Accession: S37108
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-231 <SEB>
A:Cross-references: EMBL:X74038; NTD:g398752; PID:g398753
C:Genetics:
C:Introns: 30/2
C:Superfamily: proline-rich protein

Query Match      24.5%; Score 60; DB 2; Length 231;
Best Local Similarity 50.0%; Pred. No. 12;
Matches 13; Conservative 1; Mismatches 12; Indels 0; Gaps 0.

QY      8  GDAPAGAVPGGAVPGCGPLDDGY 33
      | | | | | | | | | | | | | |
Db      152 GAAPAGGAYGSGPAGGAPAVGAY 177

Search completed: March 17, 2003, 11:38:09
Job time : 28.2549 secs

```

GenCore version 5.1.4.P5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 17, 2003, 11:32:23 ; Search time 12.6275 Seconds
(without alignments)
151.092 Million cell updates/sec

Title: US-09-855-754B-8

Perfect score: 245
Sequence: 1 GRATIRRGDAPAGAVPGCA.....PLDGMVGVDSSTVDLAQ 46

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	245	100.0	911 1	PERT_BORBR
2	232.5	94.9	922 1	PERT_BORPA
3	212	86.5	910 1	PERT_BORPE
4	64.5	26.3	565 1	HEMA_IATKP
5	64	26.1	644 1	HS71_PICAN
6	60	24.5	231 1	CUT2_CAREL
7	59.5	24.3	860 1	ELS_MOUSE
8	59	24.1	421 1	BR3A_MOUSE
9	59	24.1	423 1	BR3A_HUMAN
10	59	24.1	472 1	YAP1_MOUSE
11	59	24.1	566 1	HEMA_IADIR
12	59	24.1	1711 1	PTPO_RAT
13	58.5	23.9	128 1	MF18_MAIZE
14	58.5	23.9	480 1	RADA_MCTU
15	38	23.7	177 1	SSB_ECOLI
16	58	23.7	577 1	CST2_HUMAN
17	58	23.7	750 1	ELS_CHICK
18	57.5	23.5	529 1	SP15_STRGR
19	57.5	23.5	565 1	HEMA_IADIL
20	57.5	23.5	566 1	HEMA_IADA4
21	57.5	23.5	566 1	HEMA_IATKE
22	57.5	23.5	566 1	HEMA_IATLE
23	57.5	23.5	566 1	HEMA_IAPUE
24	57.5	23.5	566 1	HEMA_IASH2
25	57.5	23.5	566 1	HEMA_IASHS
26	57.5	23.5	566 1	HEMA_IASIN
27	57.5	23.5	566 1	HEMA_IATNU
28	57.5	23.5	566 1	HEMA_IATNU
29	57	23.3	566 1	HEMA_IATNU
30	57	23.3	759 1	MAO2_ECOLI
31	57	23.3	759 1	MAO2_SALTY
32	57	23.3	1772 1	MSPI_PLAYO
33	56.5	23.1	428 1	NSDI_PSRSE

34	56.5	23.1	547 1	HEMA_IATKP	P07976 Influenza a
35	56.5	23.1	547 1	HEMA_IATKP	P07977 Influenza a
36	56.5	23.1	562 1	HEMA_IADAP	P03451 Influenza a
37	56.5	23.1	564 1	HEMA_IATKS	P03452 Influenza a
38	56.5	23.1	564 1	HEMA_IAMAP	P87506 Influenza a
39	56.5	23.1	574 1	LOL1_HUMAN	P08397 homo sapien
40	56.5	23.1	855 1	GCFC_MOUSE	P58501 mus muscucu
41	56	22.9	172 1	YD52_SCHPO	Q10307 schizosacch
42	56	22.9	367 1	MORV_MYCLE	Q9CB48 mycobacteri
43	56	22.9	864 1	ELS_RAT	Q99372 rattus norv
44	55.5	22.7	416 1	PGK_TIRRE	P14228 trichoderma
45	55.5	22.7	565 1	HEMA_IADL1	P04661 Influenza a

ALIGNMENTS

RESULT 1	ID	PERT_BORBR	STANDARD	PRT	911 AA.
AC	003035	PERT_BORBR	STANDARD	PRT	911 AA.
DT	01-JUL-1993	(Rel. 26, Created)			
DT	01-JUL-1993	(Rel. 26, Last sequence update)			
DT	15-JUN-2002	(Rel. 41, Last annotation update)			
DE	Pertactin precursor (Outer membrane protein P.68) (P.94).				
OS	Bordetella bronchiseptica (Alcaligenes bronchisepticus).				
OC	Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;				
OC	Bordetella.				
OX	NCBI_TaxID=518;				
RN	[1]				
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 35-44.				
RC	STRAIN=CN7531;				
RA	MEDLINE=92407514; PubMed=1527510;				
RA	Li J.L., Fairweather N.F., Novotny P., Dougan G., Charles I.G.;				
RT	"Cloning, nucleotide sequence and heterologous expression of the				
RT	protective outer-membrane protein P.68 pertactin from Bordetella				
RT	bronchiseptica";				
RL	J. Gen. Microbiol. 138:1697-1705(1992).				
CC	-1- FUNCTION: AGGLOTTININ THAT BINDS TO EUKARYOTIC CELLS; A PROCESS				
CC	MEDIATED BY THE R-G-D SEQUENCE. PERTACTIN MAY HAVE A ROLE IN				
CC	BACTERIAL ADHESION, AND THUS PLAY A ROLE IN VIRULENCE.				
CC	-1- SUBUNIT: MONOMER.				
CC	-1- SUBCELLULAR LOCATION: Outer membrane.				
CC	-1- DISEASE: MAY CONTRIBUTE TO THE DISEASE STATE OF WHOOPING COUGH.				
CC	-1- MISCELLANEOUS: SYNTHESIZED ONLY IN THE PRESENCE OF LOW MG(2+)				
CC	CONCENTRATIONS.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@sib-sib.ch).				
CC	-----				
DR	EMBL: X54815; CA38584.1; -				
DR	EMBL: A19180; CA01453.1; -				
DR	PIR: A47675; A47675.				
DR	InterPro: IPR004899; Pertactin-sup.				
DR	InterPro: IPR003992; Pertactin.				
DR	InterPro: IPR003991; Pertactin_vir.				
DR	Pfam: PF03212; Pertactin; 1.				
DR	PRINTS: PR01482; PERTACTIN.				
DR	PRINTS: PR01484; PERTACTIN.FAMILY.				
KW	Outer membrane; Signal; Virulence; Repeat.				
FT	SIGNAL	1	34		
FT	CHAIN	35	911		P.94.
FT	CHAIN	35	712		PERTACTIN (P.68).
FT	PROPEP	713	911		POTENTIAL.
FT	SITE	260	262		CELL ATTACHMENT SITE (INVOLVED IN
FT					ADHESION TO VARIOUS EUKARYOTIC CELL
FT					LINES).

FT SITE 701 703 CELL ATTACHMENT SITE (POTENTIAL).
 FT DOMAIN 266 275 3 X 5 AA TANDEM REPEATS OF G-G-A-V-P.
 FT REPEAT 266 270 1.
 FT REPEAT 271 275 2.
 FT REPEAT 276 280 3 (APPROXIMATE).
 FT DOMAIN 570 601 7 X 3 AA REPEATS OF P-Q-P.
 SQ SEQUENCE 911 AA; 93995 MW; 3078DF6EC2D987A1 CRC64;
 Query Match 100.0%; Score 245; DB 1; Length 911;
 Best Local Similarity 100.0%; Pred. No. 1, 1e-19;
 Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 254 QRAIRRGDAPAGAVPGAVGFGFPLDGMVGVDSSTVDLAQ 299
 QY 1 QRAIRRGDAPAGAVPGAVGFGFPLDGMVGVDSSTVDLAQ 46
 |||||||
 DB 254 QRAIRRGDAPAGAVPGAVGFGFPLDGMVGVDSSTVDLAQ 299
 |||||||
 RESULT 2
 PERT_BORPA STANDARD; PRT; 922 AA.
 ID PERT_BORPA
 AC P24328;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Pertactin precursor (Outer membrane protein P.70) (P.95).
 GN PRN.
 OS Bordetella parapertussis.
 OC Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;
 OC Bordetella.
 OX NCBI_Taxid=519;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CN2591;
 RX MEDLINE=91251771; PubMed=2041476;
 RA Li J., Dougan G., Novotny P., Charles I.G.;
 RT "P.70 pertactin, an outer-membrane protein from Bordetella
 parapertussis: cloning, nucleotide sequence and surface expression in
 Escherichia coli.";
 RL Mol. Microbiol. 5:409-417(1991).
 CC -1- FUNCTION: AGGLUTINOGEN THAT BINDS TO EUKARYOTIC CELLS; A PROCESS
 MEDIATED BY THE R-G-D SEQUENCE. PERTACTIN MAY HAVE A ROLE IN
 BACTERIAL ADHESION, AND THUS PLAY A ROLE IN VIRULENCE.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SUBCELLULAR LOCATION: Outer membrane.
 CC -1- DISEASE: MAY CONTRIBUTE TO THE DISEASE STATE OF WHOOPING COUGH.
 CC -1- MISCELLANEOUS: SYNTHESIZED ONLY IN THE PRESENCE OF LOW MG(2+)
 CONCENTRATIONS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X54547; CAA38419.1;
 DR EMBL: A26124; CAA01786.1;
 DR EMBL: A19182; CAA01454.1;
 DR PIR: S15204; S15204.
 DR PIR: S14659; S14659.
 DR InterPro: IPR004899; Pertactin_sup.
 DR InterPro: IPR003992; pertactin.
 DR InterPro: IPR003991; pertactin_vir.
 DR Pfam: PF03212; Pertactin; 1.
 DR PRINTS: PR01482; PERTACTIN.
 DR PRINTS: PR01484; PERTACTINFAMILY.
 KM Outer membrane; signal; Virulence; Repeat.
 FT SIGNAL 1 34
 FT CHAIN 35 922 P.95. POTENTIAL.
 FT CHAIN 35 647 PERTACTIN (P.70).
 FT PROPEP 648 922
 FT SITE 260 262 CELL ATTACHMENT SITE (INVOLVED IN

FT SITE 266 290 ADHESION TO VARIOUS EUKARYOTIC CELL
 FT DOMAIN 266 270 LINES).
 FT REPEAT 266 270 4 X 5 AA TANDEM REPEATS OF G-G-A-V-P.
 FT REPEAT 271 275 1.
 FT REPEAT 276 280 2.
 FT REPEAT 281 285 3.
 FT REPEAT 575 603 4 (APPROXIMATE).
 SQ SEQUENCE 922 AA; 95178 MW; 3DF7BC58DA712478 CRC64;
 Query Match 94.9%; Score 232.5; DB 1; Length 922;
 Best Local Similarity 90.2%; Pred. No. 2, 6e-18;
 Matches 46; Conservative 0; Mismatches 0; Indels 5; Gaps 1;
 Db 254 QRAIRRGDAPAGAVPGAVGFGFPLDGMVGVDSSTVDLAQ 304
 QY 1 QRAIRRGDAPAGAVPGAVGFGFPLDGMVGVDSSTVDLAQ 46
 |||||||
 DB 254 QRAIRRGDAPAGAVPGAVGFGFPLDGMVGVDSSTVDLAQ 304
 |||||||
 RESULT 3
 PERT_BORPE STANDARD; PRT; 910 AA.
 ID PERT_BORPE
 AC P14283;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Pertactin precursor (Outer membrane protein P.69) (P.93).
 GN PRN OR OMP69A.
 OS Bordetella pertussis.
 OC Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;
 OC Bordetella.
 OX NCBI_Taxid=520;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC STRAIN=CN2892;
 RX MEDLINE=89264462; PubMed=2542937;
 RA Charles I.G., Dougan G., Pickard D., Chatfield S., Smith M.,
 RT Novotny P., Morrissey P., Fairweather N.F.;
 RT "Molecular cloning and characterization of protective outer membrane
 protein P.69 from Bordetella pertussis.";
 RL Proc. Natl. Acad. Sci. U.S.A. 86:3554-3558(1989).
 CC [2]
 CC REVISIONS TO 264 AND 332.
 CC MEDLINE=92407514; PubMed=1527510;
 CC Li J.L., Fairweather N.F., Novotny P., Dougan G., Charles I.G.;
 RT "Cloning, nucleotide sequence and heterologous expression of the
 protective outer-membrane protein P.68 pertactin from Bordetella
 bronchiseptica.";
 RL J. gen. Microbiol. 138:1697-1705(1992).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
 RX MEDLINE=96196517; PubMed=8609998;
 RA Emsley P., Charles I.G., Fairweather N.F., Isaacs N.W.;
 RT "Structure of Bordetella pertussis virulence factor P.69 pertactin.";
 RL Nature 381:90-92(1996).
 CC -1- FUNCTION: AGGLUTINOGEN THAT BINDS TO EUKARYOTIC CELLS; A PROCESS
 MEDIATED BY THE R-G-D SEQUENCE. PERTACTIN MAY HAVE A ROLE IN
 BACTERIAL ADHESION, AND THUS PLAY A ROLE IN VIRULENCE.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SUBCELLULAR LOCATION: Outer membrane.
 CC -1- DISEASE: MAY CONTRIBUTE TO THE DISEASE STATE OF WHOOPING COUGH.
 CC -1- MISCELLANEOUS: SYNTHESIZED ONLY IN THE PRESENCE OF LOW MG(2+)
 CONCENTRATION.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: J04560; MAA22980.1; AL1_SEQ.


```

DR   PIR: A32560; A32560.      pertactin_sup.
DR   InterPro: IPR004899;      pertactin.
DR   InterPro: IPR003992;      pertactin_vir.
DR   Pfam: PF03212; Pertactin; 1.
DR   PRINTS: PR01482; PERTACTIN.
DR   PRINTS: PR01484; PRYCTAFAMILY.
KW   Outer membrane; Signal; Virulence; Repeat.
FT   SIGNAL                     1       34
FT   CHAIN                      35       910    P. 93.
FT   PROPEP                     712     910    POTENTIAL.
FT   SITE                        260          CELL ATTACHMENT SITE (INVOLVED IN
                                           ADHESION TO VARIOUS EUKARYOTIC CELL
                                           LINES).
FT   DOMAIN                     266       290    5 X 5 AA TANDEM REPEATS OF G-G-A-V-P.
FT   REPEAT                     266       270    1.
FT   REPEAT                     271       275    2.
FT   REPEAT                     276       280    3.
FT   REPEAT                     281       285    4 (APPROXIMATE).
FT   REPEAT                     286       290    5 (APPROXIMATE).
FT   DOMAIN                     579        593    5 X 3 AA TANDEM REPEATS OF P-Q-P.
SQ   SEQUENCE                  910 AA: 93452 MW: A169871E20A227DB CRC64.

Query Match                               86.5%; Score 212; DB 1; Length 910;
Best Local Similarity                    75.0%; Pred. No. 4.7e-16;
Matches 42; Conservative 3; Mismatches 1; Indels 10; Gaps 1.

OY   1 ORATTTRRGDPAAGGAVPGGAV-----PGGFGPLIDGGKYGVDSYSTVDLQ 46
Db    254 ORATTTRRGDPAAGGAVPGGAVPGGFGPGGFGPVLDGWVGDVSGSSVELAQ 309
      |||.....|||.....|.....|.....|.....|.....|.....|.....|
RESULT 4
ID    HEMA_IATKP                STANDARD:      PRT:      565 AA.
AC    P03456.
DT    21-JUL-1986 (Rel. 01, Created)
DT    01-AUG-1991 (Rel. 19, Last sequence update)
DT    16-OCT-2001 (Rel. 40, Last annotation update)
DE    Hemagglutinin precursor [Contains: Hemagglutinin HA1 chain;
GN     Hemagglutinin HA2 chain].
HA.
OS    Influenza A virus (strain A/Turkey/Ontario/6/118/99).
OC    Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
CC    Influenza A viruses; Influenzavirus A.
OX    NCBI_TaxID=11471;
RN     [1]
RP     SEQUENCE FROM N.A.
RX     MEDLINE=91220697; PubMed=2024485;
RA     Nobusawa E., Aoyama T., Kato H., Suzuki Y., Tateo Y., Nakajima K.;
RT     "Comparison of complete amino acid sequences and receptor-binding
RT     properties among 13 serotypes of hemagglutinins of Influenza A
RT     viruses."
RL     Virology 182:475-485(1991).
RN     [2]
RX     SEQUENCE OF 1-99 FROM N.A.
RA     MEDLINE=82150925; PubMed=6174976;
RA     Alt G.M.;
RT     "Sequence relationships among the hemagglutinin genes of 12 subtypes
RT     of Influenza A virus."
RL     Proc. Natl. Acad. Sci. U.S.A. 78:7639-7643(1981).
CC    -1- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC    CELL RECEPTORS AND FOR INITIATING INFECTION.
CC    -1- SUBUNIT: HOMOTRIMER, EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC    (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
CC    -1- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
CC    -----
CC    This SWISS-PROT entry is copyright. It is produced through a collaboration
CC    between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC    the European Bioinformatics Institute. There are no restrictions on its
CC    use by non-profit institutions as long as its content is in no way
CC    modified and this statement is not removed. Usage by and for commercial

```

```
CC CC entitles requires a license agreement (See http://www.isb-sib.ch/announce/  
or send an email to license@sib-sib.ch).  
-----  
DR EMBL D90304; BAAL4334.1; -;  
DR EMBL J02089; AAA43177.1; -;  
DR PIR F39987; HMITVN.  
DR PIR C33157; C33157.  
DR PIR A04067; A04067.  
DR HSSP P03437; 1HTM.  
DR InterPro: IPR001364; Hemagglutn.  
DR Pfam: PF00509; Hemagglutin1.  
DR PRINTS: PR00329; HEMAGGLUTIN12.  
DR Prodom: PD000225; Hemagglutin; 1.  
KV Envelope protein; Hemagglutinin; Glycoprotein; Signal.  
FT SIGNAL 1 17  
FT CHAIN 18 342 HEMAGGLUTININ HA1 CHAIN.  
FT CHAIN 344 565 HEMAGGLUTININ HA2 CHAIN.  
FT CARBOHYD 26 26 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 27 27 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 139 139 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 150 150 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 222 222 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 303 303 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 310 310 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 497 497 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 524 524 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 556 556 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CONFLICT 9 9 T -> ML (IN REF. 2).  
FT CONFLICT 35 36 I -> T (IN REF. 2).  
FT CONFLICT 89 90 DO -> VN (IN REF. 2).  
SQ SEQUENCE 565 AA; 63518 MW; B8C2B57B63A2E5C2 CRC64;  
  
Query Match 26.3%; Score 64.5; DB 1; Length 565;  
Best Local Similarity 43.8%; Pred. No. 5.5;  
Matches 14; Conservative 4; Mismatches 9; Indels 3; Gaps 2;  
  
OY 11 PAG--GAVPGCAVPGCGFPLLDGTGVDSVS 40  
| | | : | : | : | : ||||| 1:1  
DB 342 PRLGLGAI-AGFIEGGWSGMIDGWTFGHHSNS 372  
  
RESULT 5  
HS71_PICAN STANDARD: PROT: 644 AA.  
ID HS71_PICAN  
AC P53421;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Heat-shock protein 70 1 (HSP72).  
GN HSA1.  
OS Pichia angusta (Yeast) (Hansenula polymorpha).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Pichia.  
OX NCBI_TaxId=4905;  
[1]  
SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RP STRAIN=ATCC 34438;  
RC MEDLINE=96437974; PubMed=8840502;  
RX Tloraenko V.I., Evers M.E., Diesel A., Samyn B., van Beeumen J.,  
RA Roggenkamp R.O., Kiel J.A.K.W., van der Klei I.J., Veenhuis M.;  
RT "Identification and characterization of cyclophilin Hansenula  
polymorpha proteins belonging to the Hsp70 protein family.*";  
RL Yeast 12:849-857(1996).  
CC -!- FUNCTION: ACTS AS A CHAPERONE (BY SIMILARITY).  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -!- INDUCTION: BY HEAT SHOCK.  
CC -!- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial
```

FT	REPEAT	79	82	2.				
FT	REPEAT	90	93	3.				
FT	REPEAT	105	108	4.				
FT	REPEAT	114	117	5.				
FT	REPEAT	121	124	6.				
FT	REPEAT	137	140	7.				
FT	REPEAT	153	156	8.				
FT	REPEAT	169	172	9.				
FT	REPEAT	192	195	10.				
FT	REPEAT	208	211	11.				
FT	REPEAT	218	221	12.				
SEQ	SEQUENCE	231 AA;	20833 MW;	A48C9C5496DBE797 CRC64;				
Query Match								
Best Local Similarity		24.5%;	Score 60;	DB 1;				
Matches 13;		Conservative 1;	Mismatches 12;	Indels 0;				
			Gaps	0				
QY	8 GDAPAGAVPGAVPGFGLIDGWY 33							
	1 111111 1 11 1 1 1							
Db	152 GAAPAGAVQSGAPFGAAPAVGAY 177							
RESULT 7								
ELS_MOUSE								
ID	ELS_MOUSE	STANDARD;	PRT;	860 AA.				
AC	P54320;							
DT	01-OCT-1996 (Rel. 34, Created)							
DT	01-OCT-1996 (Rel. 34, Last sequence update)							
DT	01-OCT-1996 (Rel. 34, Last annotation update)							
DE	Elastin precursor (Tropoelastin).							
ELN								
OS	Mus musculus (Mouse).							
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;							
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.							
NCBI	taxid=10090;							
RN	[1]							
RP	SEQUENCE FROM N.A.							
RC	STRAIN-BALB/c; TISSUE=Lung;							
RA	MEDICAL-95130069; PubMed=7829060;							
RA	Wyder K.S., Sechler J.L., Boyd C.D., Passmore H.C.;							
RT	"Use of an intron polymorphism to localize the tropoelastin gene to							
RT	mouse chromosome 5 in a region of linkage conservation with human							
RT	chromosome 7.;"							
RL	Genomics 23:125-131(1994).							
CC	-1- FUNCTION: MAJOR STRUCTURAL PROTEIN OF TISSUES SUCH AS AORTA AND							
CC	NUCHAL LIGAMENT, WHICH MUST EXPAND RAPIDLY AND RECOVER COMPLETELY.							
CC	-1- SUBUNIT: THE POLYMERIC ELASTIN CHAINS ARE CROSS-LINKED TOGETHER							
CC	INTO AN EXTENSIBLE 3D NETWORK.							
CC	-1- SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX OF ELASTIC FIBERS.							
CC	-1- PTM: THE CROSSLINKS ARE MADE OF DEAMINATED LYS.							
CC	-----							
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration							
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -							
CC	the European Bioinformatics Institute. There are no restrictions on its							
CC	use by non-profit institutions as long as its content is in no way							
CC	modified and this statement is not removed. Usage by and for commercial							
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/							
CC	or send an email to license@isb-sib.ch).							
CC	-----							
DR	EMBL; U08210; AAA80155.1; -							
DR	MGI; MG1:95317; ELP.							
DR	InterPro; IPR003979; tropoelastin.							
DR	PRINTS; PRO1500; TROPOELASTIN.							
KW	Structural protein; Repeat; Signal; Connective tissue.							
FT	SIGNAL	1	27	POTENTIAL.				
FT	CHAIN	28	860	ELASTIN				
SEQ	SEQUENCE	860 AA;	71955 MW;	OC0B5A81ED7F1 CRC64;				
Query Match								
Best Local Similarity		24.3%;	Score 59.5;	DB 1;				
Matches 14;		Conservative 2;	Mismatches 10;	Indels 13;				
			Gaps	1;				

between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL: U10063; AAA57161.1; -
 DR EMBL: U10062; AAA57161.1; JOINED.
 DR EMBL: L20433; AAA65605.1; -
 DR EMBL: X64624; CAA45907.1; -
 DR HSSP: P10037; LAU7.
 DR TRANSFAC: T01876; -
 DR TRANSFAC: T04463; -
 DR Genew; HGNC:9218; POU4F1.
 DR MIM: 601632; -
 DR InterPro; IPR001356; Homeobox.
 DR InterPro; IPR000327; POU_domain.
 DR Pfam; PF00046; homeobox; 1.
 DR Pfam; PF00157; pou; 1.
 DR PRINTS: PR00028; POUDOMAIN.
 DR ProDom; PD000010; Homeobox; 1.
 DR ProDom; PD000583; POU_domain; 1.
 DR SMART; SM00389; HOX; 1.
 DR SMART; SM00352; POU; 1.
 DR PROSITE; PS00027; HOMEBOX_1; 1.
 DR PROSITE; PS00035; POU_1; 1.
 DR PROSITE; PS00465; POU_2; 1.
 DR PROSITE; PS50071; HOMEBOX_2; 1.
 DR Transcription regulation; Nuclear protein; DNA-binding; Homeobox; Developmental protein.
 KW DOMAIN 57
 FT DOMAIN 100 108 POU-IV BOX.
 FT 130 189 POLY-HIS.
 FT DOMAIN 130 189 GLY-RICH.
 FT 207 260 ALA-RICH.
 FT DOMAIN 267 341 POU.
 FT 359 418 HOMEBOX.
 FT DNA_BIND 91 92 TS -> H (IN REF. 3).
 FT CONFLICT 99 99 A -> R (IN REF. 1).
 FT CONFLICT 133 135 GAG -> ARR (IN REF. 3).
 FT CONFLICT 135 137 MISSING (IN REF. 2).
 FT CONFLICT 139 142 MISSING (IN REF. 3).
 FT CONFLICT 149 150 GP -> AA (IN REF. 3).
 FT CONFLICT 155 156 GP -> PR (IN REF. 3).
 FT CONFLICT 174 175 GP -> AA (IN REF. 3).
 FT CONFLICT 189 189 G -> A (IN REF. 3).
 FT CONFLICT 240 240 A -> S (IN REF. 2).
 FT CONFLICT 252 252 MISSING (IN REF. 3).
 FT CONFLICT 256 256 A -> R (IN REF. 3).
 FT CONFLICT 276 278 AER -> GS (IN REF. 3).
 FT CONFLICT 343 345 AOR -> PS (IN REF. 3).
 SQ SEQUENCE 423 AA; 42939 MW; B3AE4732E1309F34 CRC64;

Query Match 24.1%; Score 59; DB 1; Length 423;
 Best Local Similarity 57.9%; Pred. No. 17;
 Matches 11; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 8 GDAPGAGVPGGAVPGGFG 26
 DB 153 GGGPGGGGPGGGPGGG 171

RESULT 10
 YAPL_MOUSE STANDARD; PRT; 472 AA.
 AC P46938;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-JUN-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE 65 kDa yes-associated protein (YAP65).
 GN YAP1 OR YAP65 OR YAP.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NIH Swiss; TISSUE=Embryo;
 RX MEDLINE=95301570; PubMed=7782338;
 RA Sudol M., Bork P., Einbond A., Kastury K., Druck T., Negrini M., Huebner K., Lehman D.;
 RT "Characterization of the mammalian YAP (yes-associated protein) gene and its role in defining a novel protein module, the WW domain.";
 RL J. Biol. Chem. 270:14733-14741(1995).
 CC -1- FUNCTION: BINDS TO THE SH3 DOMAIN OF THE YES KINASE (BY SIMILARITY).
 CC -1- PTM: PHOSPHORYLATED ON SERINE RESIDUES (BY SIMILARITY).
 CC -1- SIMILARITY: CONTAINS 2 WW DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL: X80508; CAA56673.1; -
 DR MGD; MGI:103262; Yap.
 DR InterPro; IPR002349; WW.
 DR InterPro; IPR001202; WW_Rsp5_WWP.
 DR Pfam; PF00397; WW; 2.
 DR PRINTS; PR00403; WWDOMAIN.
 DR SMART; SM00456; WW; 2.
 DR PROSITE; PS01159; WW_DOMAIN_1; 2.
 DR PROSITE; PS50020; WW_DOMAIN_2; 2.
 DR Phosphorylation; Repeat.
 FT DOMAIN 156 189 WW 1.
 FT 215 248 WW 2.
 SQ SEQUENCE 472 AA; 50703 MW; 174FD33E0381126C CRC64;

Query Match 24.1%; Score 59; DB 1; Length 472;
 Best Local Similarity 42.4%; Pred. No. 19;
 Matches 14; Conservative 1; Mismatches 18; Indels 0; Gaps 0;

QY 10 APAGAVPGGAVPGGPGPLDGVGVDSSTV 42
 DB 199 APASPAVPQTLNMSASGPLEDGEQAMTQDGEV 231

RESULT 11
 HEMA_IADIR STANDARD; PRT; 566 AA.
 AC P11132;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hemagglutinin precursor [Contains: Hemagglutinin HA1 chain; Hemagglutinin HA2 chain].
 GN HA.
 OS Influenza A virus (strain A/Duck/Ireland/113/83).
 CC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
 CC Influenza A viruses; Influenzavirus A.
 CX NCBI_TaxID=11365;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87207664; PubMed=3576972;
 RA Kawoka Y., Nestorowicz A., Alexander D.J., Webster R.G.;
 RT "Molecular analyses of the hemagglutinin genes of H5 Influenza viruses: origin of a virulent turkey strain.";
 RL Virology 158:218-227(1987).
 CC -1- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO CELL RECEPTORS AND FOR INITIATING INFECTION.
 CC -1- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
 CC -1- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.

FT CARBOHYD 882 882 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 970 970 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 982 982 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 1711 AA; 187292 MW; 0F042DD1A47A18A0 CRC64;

Query Match 24.1%; Score 59; DB 1; Length 1711;
 Best Local Similarity 38.0%; Pred. No. 63;
 Matches 19; Conservative 8; Mismatches 19; Indels 4; Gaps 2;

QY 1 QARATIRGDAP---AGGAVGAVGPGGGLLGG-WGCVVDSYDVLQ 46
 DB 416 RRLVLSDDAPGSLGNISVSGATHYFGLVFGAHRVDIASSTGDISQ 465

RESULT 13
 MF18_MAIZE
 ID MF18_MAIZE STANDARD; PRT; 128 AA.
 AC P32439;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE MFS18 protein precursor.
 GN MFS18.
 OS Zea mays (Maize).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
 OC Panicoideae; Andropogoneae; Zea.
 NCBI_Taxid=4577;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=CV. BELO; TISSUE=Tassel;
 RX MEDLINE=94004987; PubMed=8401606;
 RA Wright S.Y., Suter M.-M., Bell P.J., Vaudin M., Greenland A.J.;
 RT "Isolation and characterization of male flower cDNAs from maize."
 RL Plant J. 3:41-49(1993).
 CC -1- TISSUE SPECIFICITY: ENHANCED EXPRESSION IN MALE FLOWERS.
 CC ACCUMULATES IN THE GLUMES AND IN ANTER WALLS, PALEAS AND LEMMAS
 CC OF MATURE FLORETS.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED THROUGHOUT TASSEL GROWTH UP UNTIL
 CC MATURE POLLEN IS PRODUCED IN THE ANTHERS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

CC EMBL; X67324; CAA47738.1;
 CC DR PIR; S25103; S25103.
 CC DR MAIRED; 78601; -

FT SIGNAL Repeat: 1 25 POTENTIAL.
 FT CHAIN 26 128 MFS18 PROTEIN.
 FT DOMAIN 64 92 3 X APPROXIMATE TANDEM REPEATS.
 FT REPEAT 69 80 1-1.
 FT REPEAT 81 92 1-2.
 FT REPEAT 81 92 1-3.
 FT DOMAIN 64 116 8 X 4 AA APPROXIMATE REPEATS.
 FT REPEAT 64 67 2-1.
 FT REPEAT 69 72 2-2.
 FT REPEAT 74 77 2-3.
 FT REPEAT 79 82 2-4.
 FT REPEAT 86 89 2-5.
 FT REPEAT 91 94 2-6.
 FT REPEAT 104 107 2-7.
 FT REPEAT 113 116 2-8.
 SQ SEQUENCE 128 AA; 12535 MW; 7661ECC596E0D778 CRC64;

Query Match 23.9%; Score 58.5; DB 1; Length 128;
 Best Local Similarity 38.2%; Pred. No. 6.1;
 Matches 13; Conservative 5; Mismatches 7; Indels 9; Gaps 1;

QY 8 GDAPAGAVPGAVPG-----GGPILGGW 32
 DB 55 GASPAFGGLPGSGIPGSSIPGFSMPGSGSLPGF 88

RESULT 14
 RADA_MYCTU
 ID RADA_MYCTU STANDARD; PRT; 480 AA.
 AC 035370;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE DNA repair protein rada homolog (DNA repair protein sms homolog).
 GN RADA OR RV3585 OR MT3691 OR MYV024.03.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
 OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
 NCBI_Taxid=1773;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=H37RV;
 RX MEDLINE=98293987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eigmler K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sultson J.E., Taylor K., Whitehead S., Barrett B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence."
 RL Nature 393:537-544(1998).
 RN [2]

RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ernolova M.D., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bissel W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains."
 RL Submitted (APR-2001) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: MAY PLAY A ROLE IN THE REPAIR OF ENDOGENOUS ALKYLATION
 CC DAMAGE (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE RECA FAMILY. RADA SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

CC EMBL; AL022075; CAA17854.1;
 CC DR EMBL; AE007169; AAK48049.1; -
 CC DR TIGR; MT3691; -
 CC DR Tuberculist; RV3585; -
 CC DR InterPro; IPR003593; AAA_ATPase.
 CC DR InterPro; IPR001984; Ion_endopep.
 CC DR InterPro; IPR001553; RECA.
 CC DR InterPro; IPR004504; SMS.
 CC DR PRINTS; PR00830; ENDOLAPLASE.
 CC DR PRODOM; PD000229; RECA.1.
 CC DR SMART; SMO0382; AAA.1.
 CC DR TIGRFRAMS; TIGR00416; sms; 1.
 CC DR PROSITE; PS0162; RECA.2; 1.
 CC DNA repair; ATP-binding; Zinc-finger; DNA-binding; Complete proteome.
 CC C4-type (POTENTIAL).
 FT NP_BIND 95 102 ATP (POTENTIAL).

SO SEQUENCE 480 AA; 49913 MW; DDECEA049F7099BC CRC64;
Query Match 23.9%; Score 58.5; DB 1; Length 480;
Best Local Similarity 33.3%; Pred. No. 21;
Matches 22; Conservative 3; Mismatches 16; Indels 25; Gaps 2;
OY 4 TIRRGDAPAGVAVP-----GGAVPGGFGFLDQWGVVDVSD 39
DB 45 TRRRVAAASAGVAVPISAVDAHTRPCPTGIDELDNVLGGIVPGSV-TLLAGDPGVGKST 103
OY 40 STVDIA 45
DB 104 LLEEVA 109
RESULT 15
SSB_ECOLI STANDARD; PRT; 177 AA.
ID SSB_ECOLI
AC P02339;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Single-strand binding protein (SSB) (Helix-destabilizing protein).
GN SSB OR EXRB OR LEXC OR B4059 OR Z5658 OR ECS5041.
OS Escherichia coli, and
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia
CX NCBI_TaxID=562, 83334;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-52.
RX MEDLINE=82037821; PubMed=6270666;
RA Sancar A., Williams K.R., Chase J.W., Rupp W.D.;
RT "Sequences of the ssb gene and protein."
RL Proc. Natl. Acad. Sci. U.S.A. 78:4274-4278(1981).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=K12 / MG1655;
RA MEDLINE=94089392; PubMed=8265357;
RA Blatter F.R., Burland V.D., Plunkett G. III, Sofia H.J.;
RL Daniels D.L.;
RT "Analysis of the Escherichia coli genome. IV. DNA sequence of the
RT region from 89.2 to 92.8 minutes."
RL Nucleic Acids Res. 21:5408-5417(1993).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=O157:H7 / EDL933 / ATCC 700927;
RA MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobbeck E.J., Davis N.W., Lam A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7."
RL Nature 409:529-533(2001).
RN [4]
RP SEQUENCE FROM N.A.
RX STRAIN=O157:H7 / RIMD 0509952;
RA MEDLINE=21156231; PubMed=11258796;
RA Hayaishi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsudo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasaki C., Ogasawara N., Yasunaga T.,
RA Kubara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12."
RL DNA Res. 8:11-22(2001).
RN [5]
RP SEQUENCE OF 1-40.
RX MEDLINE=82186711; PubMed=7042342;
RA Beyreuther K., Berthold-Schmidt V., Gelder K.;
RT "Biological activity and a partial amino-acid sequence of Escherichia
RT coli DNA-binding protein I isolated from overproducing cells."

RL Eur. J. Biochem. 123:415-420(1982).
RN [6]
RP CHARACTERIZATION, AND REVISION TO 133.
RX MEDLINE=84111559; PubMed=6363409;
RA Chase J.W., L'Italian J.J., Murphy J.B., Spicer E.K., Williams K.R.;
RT "Characterization of the Escherichia coli SSB-113 mutant
RT single-stranded DNA-binding protein. Cloning of the gene, DNA and
RT protein sequence analysis, high pressure liquid chromatography
RT peptide mapping, and DNA-binding studies."
RL J. Biol. Chem. 259:805-814(1984).
RN [7]
RP MUTANT SSB-1.
RX MEDLINE=85006904; PubMed=6384214;
RA Williams K.R., Murphy J.B., Chase J.W.;
RT "Characterization of the structural and functional defect in the
RT Escherichia coli single-stranded DNA binding protein encoded by the
RT ssb-1 mutant gene. Expression of the ssb-1 gene under lambda pl
RT regulation."
RL J. Biol. Chem. 259:11804-11811(1984).
RN [8]
RP MUTANT SSB-1, AND DNA-BINDING.
RX MEDLINE=91108818; PubMed=1988680;
RA Bujalowski W., Lohman T.M.;
RT "Monomers of the Escherichia coli SSB-1 mutant protein bind single-
RT stranded DNA."
RL J. Mol. Biol. 217:63-74(1991).
RN [9]
RP MUTANTS SSB-200; SSB-201 AND SSB-202.
RX STRAIN=K12 / W3110;
RA MEDLINE=97284491; PubMed=9139905;
RA Reddy M., Gowrishankar J.;
RT "Identification and characterization of ssb and uup mutants with
RT increased frequency of precise excision of transposon Tn10
RT derivatives: nucleotide sequence of uup in Escherichia coli."
RL J. Bacteriol. 179:2892-2899(1997).
RN [10]
RP MOTAGENESIS, AND DNA-BINDING.
RX MEDLINE=87276593; PubMed=3301414;
RA Casas-Finet J.R., Khamsi M.I., Maki A.W., Chase J.W.;
RT "Tryptophan 54 and phenylalanine 60 are involved synergistically in
RT the binding of E. coli SSB protein to single-stranded
RT polynucleotides."
RL FEBS Lett. 220:347-352(1987).
RN [11]
RP REVIEW.
RX MEDLINE=91203755; PubMed=2087220;
RA Meyer R.R., Laine P.S.;
RT "The single-stranded DNA-binding protein of Escherichia coli."
RL Microbiol. Rev. 54:342-380(1990).
RN [12]
RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS) OF 1-135.
RX MEDLINE=97338073; PubMed=9192620;
RA Ragunathan S., Ricard C.S., Lohman T.M., Wakeman G.;
RT "Crystal structure of the homo-tetrameric DNA binding domain of
RT Escherichia coli single-stranded DNA-binding protein determined by
RT multiwavelength x-ray diffraction on the selenomethionyl protein at
RT 2.9-A resolution."
RL Proc. Natl. Acad. Sci. U.S.A. 94:6652-6657(1997).
RN [13]
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 1-135.
RX MEDLINE=20392541; PubMed=10932248;
RA Ragunathan S., Kozlov A.G., Lohman T.M., Wakeman G.;
RT "Structure of the DNA binding domain of E. coli SSB bound to ssDNA."
RL Nat. Struct. Biol. 7:648-652(2000).
CC -1- FUNCTION: THIS PROTEIN IS ESSENTIAL FOR REPLICATION OF THE
CC CHROMOSOMES AND ITS SINGLE-STRANDED DNA PHAGES. IT IS ALSO
CC INVOLVED IN DNA RECOMBINATION AND REPAIR.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SIMILARITY: BELONGS TO THE SSB FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: March 17, 2003, 11:35:38 ; Search time 46.902 Seconds
(Without alignments) 202.085 Million cell updates/sec

Title: US-09-855-754b-8

Perfect score: 245
Sequence: 1 ORATIRRGDAPAGAVPGCA.....PLLDGKYGVDVSDSTVDLAQ 46

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

1: SPTRMBL_21:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mammal:*
8: sp_mmc:*
9: sp_organelle:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	245	100.0	147	2	Q9K5H6 Bordetella
2	245	100.0	150	2	Q9K5H8 Bordetella
3	245	100.0	150	2	Q9K5H4 Bordetella
4	245	100.0	154	2	Q9K5H0 Bordetella
5	245	100.0	155	2	Q9K5H2 Bordetella
6	245	100.0	173	2	Q9ALQ5 Bordetella
7	245	100.0	178	2	Q9ALQ3 Bordetella
8	245	100.0	177	2	Q9ALQ3 Bordetella
9	245	100.0	206	2	Q9KJY2 Bordetella
10	245	100.0	911	2	Q9LAZ2 Bordetella
11	232.5	94.9	158	2	Q9K5G6 Bordetella
12	232.5	94.9	159	2	Q9K5G8 Bordetella
13	232.5	94.9	160	2	Q9K5I0 Bordetella
14	232.5	94.9	179	2	Q9AHP1 Bordetella
15	232.5	94.9	211	2	Q9KJY3 Bordetella
16	230	93.9	216	2	Q9KJY4 Bordetella

17	214.5	87.6	382	2	Q93199 Bordetella
18	214.5	87.6	905	2	Q9S6M9 Bordetella
19	214.5	87.6	905	2	Q9S6M0 Bordetella
20	212	86.5	159	2	Q9ALP4 Bordetella
21	212	86.5	160	2	Q9K5G3 Bordetella
22	212	86.5	161	2	Q9K5G2 Bordetella
23	212	86.5	387	2	Q9S3M9 Bordetella
24	212	86.5	910	2	Q9S6M1 Bordetella
25	212	86.5	910	2	Q9S2S9 Bordetella
26	210.5	85.9	181	2	Q9ALQ1 Bordetella
27	209.5	85.5	167	2	Q93705 Bordetella
28	209.5	85.5	915	2	Q9S143 Bordetella
29	208	84.9	907	2	Q9S2S7 Bordetella
30	207	84.5	851	2	Q9AIX8 Bordetella
31	207	84.5	920	2	Q93L58 Bordetella
32	205.5	83.9	912	2	Q9RSU0 Bordetella
33	204.5	83.5	184	2	Q9ALP6 Bordetella
34	204.5	83.5	195	2	Q9ALP8 Bordetella
35	201.5	82.2	905	2	Q9S3M8 Bordetella
36	76.5	31.2	1884	5	Q9NHW2
37	74.5	30.4	871	5	Q44358
38	74.5	30.4	2249	5	Q9NHW4
39	73.5	30.0	646	10	Q9M2F9
40	70	28.6	907	5	Q44359
41	69.5	28.4	383	12	Q67012
42	69.5	28.2	462	5	Q9NHW3
43	68	27.8	626	5	Q9NHW1
44	68	27.8	642	3	Q9UVM0
45	67	27.3	527	11	Q9JMC8

ALIGNMENTS

RESULT 1

ID	Q9K5H6	PRELIMINARY;	PRT;	147 AA.
AC	Q9K5H6;			
DT	01-OCT-2000 (TREMREL.15, Created)			
DT	01-OCT-2000 (TREMREL.15, Last sequence update)			
DT	01-JUN-2001 (TREMREL.17, Last annotation update)			
DE	Pertactin (P.68) (Fragment).			
GN	PRN.			
OS	Bordetella bronchiseptica (Alcaligenes bronchisepticus).			
OC	Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;			
CC	Bordetella.			
NCBI_TaxID-518;				
RP	SEQUENCE FROM N.A.			
RC	STRAIN-5;			
RX	MEDLINE-20359389; PubMed-10899896;			
RA	Boursaux-Ende C, Guiso N;			
RT	"Polymorphism of Repeated Regions of Pertactin in Bordetella			
RT	Pertussis, Bordetella pertussis, and Bordetella bronchiseptica."			
RL	Infect. Immun. 68:4815-4817(2000).			
DR	EMBL; AJ250080; CAB76434.1;			
FT	NON_TER			
FT	NON_TER			
SO	SEQUENCE	147 AA;	147	MM; 83F8E6526202E5E3 CRC64;

Query Match 100.0%; Score 245; DB 2; Length 147;
Best Local Similarity 100.0%; Pred. No. 2.1e-20;

Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ORATIRRGDAPAGAVPGAGFGLDGYGVSDSTVDLAQ 46

DB 47 ORATIRRGDAPAGAVPGAGFGLDGYGVSDSTVDLAQ 92

RESULT 2

Q9K5H8 PRELIMINARY; PRT; 150 AA.
AC Q9K5H8;

01-OCT-2000 (TREMBlrel. 15, Created)
01-OCT-2000 (TREMBlrel. 15, Last sequence update)
01-JUN-2001 (TREMBlrel. 17, Last annotation update)
Pertactin (P.68) (Fragment).
Bordetella bronchiseptica (Alcaligenes bronchisepticus).
Bacteria: Proteobacteria; beta subdivision; Alcaligenaceae;
NCBI_TaxID=518;
SEQUENCE FROM N.A.
STRAIN=LAPR:
MEDLINE=20359389; PubMed=10899896;
Boursaux-Eude C., Guiso N.;
"Polymorphism of Repeated Regions of Pertactin in Bordetella pertussis, Bordetella parapertussis, and Bordetella bronchiseptica.";
Infect. Immun. 68:4815-4817(2000).
EMBL: AJ250078; CAB76432.1; --
NON_TER 1 1
FT 150 150
SQ SEQUENCE 150 AA; 14614 MW; B1CEC7FDA16725 CRC64;
Query Match 100.0%; Score 245; DB 2; Length 150;
Best Local Similarity 100.0%; Pred. No. 2.2e-20;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 ORATIRGDAPAGGAVPGGFGFPLLDGWTGVDSSTVDLAQ 46
53 ORATIRGDAPAGGAVPGGFGFPLLDGWTGVDSSTVDLAQ 98
RESULT 3
ID 09K5H4 PRELIMINARY; PRT; 150 AA.
AC 09K5H4;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Pertactin (P.68) (Fragment).
GN PRN.
OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
OC Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;
OC Bordetella.
OX NCBI_TaxID=518;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=335;
RX MEDLINE=20359389; PubMed=10899896;
RA Boursaux-Eude C., Guiso N.;
"Polymorphism of Repeated Regions of Pertactin in Bordetella pertussis, Bordetella parapertussis, and Bordetella bronchiseptica.";
Infect. Immun. 68:4815-4817(2000).
EMBL: AJ250082; CAB76436.1; --
NON_TER 1 1
FT 150 150
SQ SEQUENCE 150 AA; 14425 MW; C3AA820B3CEAE60 CRC64;
Query Match 100.0%; Score 245; DB 2; Length 150;
Best Local Similarity 100.0%; Pred. No. 2.2e-20;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 ORATIRGDAPAGGAVPGGFGFPLLDGWTGVDSSTVDLAQ 46
50 ORATIRGDAPAGGAVPGGFGFPLLDGWTGVDSSTVDLAQ 95
RESULT 4
ID 09K5H0 PRELIMINARY; PRT; 154 AA.
AC 09K5H0;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Pertactin (P.68) (Fragment).
GN PRN.
OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
OC Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;
OC Bordetella.
OX NCBI_TaxID=518;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=335;
RX MEDLINE=20359389; PubMed=10899896;
RA Boursaux-Eude C., Guiso N.;
"Polymorphism of Repeated Regions of Pertactin in Bordetella pertussis, Bordetella parapertussis, and Bordetella bronchiseptica.";
Infect. Immun. 68:4815-4817(2000).
EMBL: AJ250084; CAB76438.1; --
NON_TER 1 1
FT 154 154
SQ SEQUENCE 154 AA; 14799 MW; 9D28B10C4249D23E CRC64;
Query Match 100.0%; Score 245; DB 2; Length 154;
Best Local Similarity 100.0%; Pred. No. 2.2e-20;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 ORATIRGDAPAGGAVPGGFGFPLLDGWTGVDSSTVDLAQ 46
53 ORATIRGDAPAGGAVPGGFGFPLLDGWTGVDSSTVDLAQ 98
RESULT 5
ID 09K5H2 PRELIMINARY; PRT; 155 AA.
AC 09K5H2;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Pertactin (P.68) (Fragment).
GN PRN.
OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
OC Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;
OC Bordetella.
OX NCBI_TaxID=518;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=335;
RX MEDLINE=20359389; PubMed=10899896;
RA Boursaux-Eude C., Guiso N.;
"Polymorphism of Repeated Regions of Pertactin in Bordetella pertussis, Bordetella parapertussis, and Bordetella bronchiseptica.";
Infect. Immun. 68:4815-4817(2000).
EMBL: AJ250084; CAB76438.1; --
NON_TER 1 1
FT 155 155
SQ SEQUENCE 155 AA; 14899 MW; E8A5C447F8CE43E CRC64;
Query Match 100.0%; Score 245; DB 2; Length 155;
Best Local Similarity 100.0%; Pred. No. 2.3e-20;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 ORATIRGDAPAGGAVPGGFGFPLLDGWTGVDSSTVDLAQ 46
54 ORATIRGDAPAGGAVPGGFGFPLLDGWTGVDSSTVDLAQ 99
RESULT 6
ID 09AL05 PRELIMINARY; PRT; 173 AA.
AC 09AL05;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Pertactin (Fragment).
GN PRN.
OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).

Pertactin (P.68) (Fragment).
GN PRN.
OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
OC Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;
OC Bordetella.
OX NCBI_TaxID=518;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=335;
RX MEDLINE=20359389; PubMed=10899896;
RA Boursaux-Eude C., Guiso N.;
"Polymorphism of Repeated Regions of Pertactin in Bordetella pertussis, Bordetella parapertussis, and Bordetella bronchiseptica.";
Infect. Immun. 68:4815-4817(2000).
EMBL: AJ250086; CAB76440.1; --
NON_TER 1 1
FT 154 154
SQ SEQUENCE 154 AA; 14799 MW; 9D28B10C4249D23E CRC64;
Query Match 100.0%; Score 245; DB 2; Length 154;
Best Local Similarity 100.0%; Pred. No. 2.2e-20;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 ORATIRGDAPAGGAVPGGFGFPLLDGWTGVDSSTVDLAQ 46
53 ORATIRGDAPAGGAVPGGFGFPLLDGWTGVDSSTVDLAQ 98
RESULT 5
ID 09K5H2 PRELIMINARY; PRT; 155 AA.
AC 09K5H2;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Pertactin (P.68) (Fragment).
GN PRN.
OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
OC Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;
OC Bordetella.
OX NCBI_TaxID=518;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=335;
RX MEDLINE=20359389; PubMed=10899896;
RA Boursaux-Eude C., Guiso N.;
"Polymorphism of Repeated Regions of Pertactin in Bordetella pertussis, Bordetella parapertussis, and Bordetella bronchiseptica.";
Infect. Immun. 68:4815-4817(2000).
EMBL: AJ250084; CAB76438.1; --
NON_TER 1 1
FT 155 155
SQ SEQUENCE 155 AA; 14899 MW; E8A5C447F8CE43E CRC64;
Query Match 100.0%; Score 245; DB 2; Length 155;
Best Local Similarity 100.0%; Pred. No. 2.3e-20;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 ORATIRGDAPAGGAVPGGFGFPLLDGWTGVDSSTVDLAQ 46
54 ORATIRGDAPAGGAVPGGFGFPLLDGWTGVDSSTVDLAQ 99
RESULT 6
ID 09AL05 PRELIMINARY; PRT; 173 AA.
AC 09AL05;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Pertactin (Fragment).
GN PRN.
OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).

OC Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;
OC Bordetella.
OX NCBI_TaxID=518;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SOI0328;
RX MEDLINE-21117018; PubMed-11179374;
RA Register K.B.;
RT "Novel Genetic and Phenotypic Heterogeneity in Bordetella
bronchiseptica pertactin.";
RL Infect. Immun. 69:1917-1921(2001).
DR EMBL; AY007262; AAG38438.1; -
DR InterPro; IPR003992; pertactin.
DR PRINTS; PRO1482; PERTACTIN.
FT NON_TER 1
FT SEQUENCE 173 AA; 16802 MW; 49132DBE8C4C81A0 CRC64;

Query Match 100.0%; Score 245; DB 2; Length 173;
Best Local Similarity 100.0%; Pred. No. 2.5e-20;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ORATIRRGDAPAGAVPGGFGPPLDGMVGVDSSTVDLAQ 46
Db 78 ORATIRRGDAPAGAVPGGFGPPLDGMVGVDSSTVDLAQ 123

RESULT 7
O9ALQ3 PRELIMINARY; PRT; 177 AA.
AC O9ALQ3;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Pertactin (Fragment).
GN PRN.
OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
OC Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;
OC Bordetella.
OX NCBI_TaxID=518;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC19395;
RX MEDLINE-21117018; PubMed-11179374;
RA Register K.B.;
RT "Novel Genetic and Phenotypic Heterogeneity in Bordetella
bronchiseptica pertactin.";
RL Infect. Immun. 69:1917-1921(2001).
DR EMBL; AY007264; AAG38440.1; -
DR InterPro; IPR003992; pertactin.
DR PRINTS; PRO1482; PERTACTIN.
FT NON_TER 1
FT SEQUENCE 177 AA; 17272 MW; 6011597CE5996ACC CRC64;

Query Match 100.0%; Score 245; DB 2; Length 177;
Best Local Similarity 100.0%; Pred. No. 2.6e-20;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ORATIRRGDAPAGAVPGGFGPPLDGMVGVDSSTVDLAQ 46
Db 82 ORATIRRGDAPAGAVPGGFGPPLDGMVGVDSSTVDLAQ 127

RESULT 8
O93ST1 PRELIMINARY; PRT; 178 AA.
AC O93ST1;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Pertactin (Fragment).
GN PRN.

OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
OC Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;
OC Bordetella.
OX NCBI_TaxID=518;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-KM22;
RX PubMed-11179374;
RA Register K.B.;
RT "Novel Genetic and Phenotypic Heterogeneity in Bordetella
bronchiseptica pertactin.";
RL Infect. Immun. 69:1917-1921(2001).
DR EMBL; AY007270; AAG38446.2; -
FT NON_TER 1
FT SEQUENCE 178 AA; 17373 MW; D42F99421BD0D066 CRC64;

Query Match 100.0%; Score 245; DB 2; Length 178;
Best Local Similarity 100.0%; Pred. No. 2.6e-20;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ORATIRRGDAPAGAVPGGFGPPLDGMVGVDSSTVDLAQ 46
Db 83 ORATIRRGDAPAGAVPGGFGPPLDGMVGVDSSTVDLAQ 128

RESULT 9
O9KJY2 PRELIMINARY; PRT; 206 AA.
AC O9KJY2;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Pertactin (Fragment).
OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
OC Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;
OC Bordetella.
OX NCBI_TaxID=518;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-16039;
RA Kell D.J.; Fenwick B.;
RT "Variability in the pertactin genes of Bordetella bronchiseptica
isolates includes the regions coding for the GGXXP and PGP families of
RT amino acid repeats.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF156770; AAF82394.1; -
FT NON_TER 1
FT SEQUENCE 206 AA; 20566 MW; 7E2435E759A97BE9 CRC64;

Query Match 100.0%; Score 245; DB 2; Length 206;
Best Local Similarity 100.0%; Pred. No. 3.1e-20;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ORATIRRGDAPAGAVPGGFGPPLDGMVGVDSSTVDLAQ 46
Db 154 ORATIRRGDAPAGAVPGGFGPPLDGMVGVDSSTVDLAQ 199

RESULT 10
O9LAE2 PRELIMINARY; PRT; 911 AA.
AC O9LAE2;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Pertactin.
GN PRN.
OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
OC Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;
OC Bordetella.
OX NCBI_TaxID=518;

```
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-B14;
RA King A., Berbers G., Hoogerhout P., Oirschot van H.F., Knipping K.,
RM Mool F.R.;
RT "role of the polymorphic region of the B.pertussis protein pertactin
in immunity: implications for the design of pertussis vaccines.";
RL Submitted (Aug-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ245927; CAB82515.1; -
DR InterPro; IPR003992; pertactin.
DR InterPro; IPR003991; pertactin_vlr.
DR InterPro; IPR004899; pertactin_sup.
DR Pfam; PF03212; Pertactin_1.
DR PRINTS; PRO1482; PERTACTIN.
DR PRINTS; PRO1484; PERTACTINFMXY.
SQ SEQUENCE 911 AA; 94093 MW; DF531A9EB438A32 CRC64;

Query Match 100.0%; Score 245; DB 2; Length 911;
Best Local Similarity 100.0%; Pred. No. 1.6e-19;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ORATIRRGDAPAGGAVPGGAVPGFGLDGMVGVDSSTVDLAQ 46
Db 254 ORATIRRGDAPAGGAVPGGAVPGFGLDGMVGVDSSTVDLAQ 299

RESULT 11
ID 09K5G6 PRELIMINARY; PRT; 158 AA.
AC 09K5G6;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Pertactin (P.68) (Fragment).
GN PRN.
OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
OC Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;
OC Bordetella.
OX NCBI_TaxID=518;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CAT1;
RX MEDLINE-20359389; PubMed-10899896;
RA Boursaux-Eude C., Guiso N.;
RT "Polymorphism of Repeated Regions of Pertactin in Bordetella
pertussis, Bordetella parapertussis, and Bordetella bronchiseptica.";
RL Infect. Immun. 68:4815-4817(2000).
DR EMBL; AJ250090; CAB76444.1; -
FT NON_TER 1
FT NON_TER 158
SQ SEQUENCE 158 AA; 15062 MW; F8ECD81013EBFDC CRC64;

Query Match 94.9%; Score 232.5; DB 2; Length 158;
Best Local Similarity 90.2%; Pred. No. 6e-19;
Matches 46; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

OY 1 ORATIRRGDAPAGGAVPGGAVPGFGLDGMVGVDSSTVDLAQ 46
Db 54 ORATIRRGDAPAGGAVPGGAVPGFGLDGMVGVDSSTVDLAQ 104

RESULT 12
ID 09K5G8 PRELIMINARY; PRT; 159 AA.
AC 09K5G8;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Pertactin (P.68) (Fragment).
GN PRN.
OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
OC Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;
OC Bordetella.
OX NCBI_TaxID=518;
RN [1]
RP SEQUENCE FROM N.A.
```

```
OX NCBI_TaxID=518;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-DEL;
RX MEDLINE-20359389; PubMed-10899896;
RA Boursaux-Eude C., Guiso N.;
RT "Polymorphism of Repeated Regions of Pertactin in Bordetella
pertussis, Bordetella parapertussis, and Bordetella bronchiseptica.";
RL Infect. Immun. 68:4815-4817(2000).
DR EMBL; AJ250088; CAB76442.1; -
FT NON_TER 1
FT NON_TER 159
SQ SEQUENCE 159 AA; 15181 MW; 9F315BB478659E95 CRC64;

Query Match 94.9%; Score 232.5; DB 2; Length 159;
Best Local Similarity 90.2%; Pred. No. 6.1e-19;
Matches 46; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

OY 1 ORATIRRGDAPAGGAVPGGAVPGFGLDGMVGVDSSTVDLAQ 46
Db 53 ORATIRRGDAPAGGAVPGGAVPGFGLDGMVGVDSSTVDLAQ 103

RESULT 13
ID 09K510 PRELIMINARY; PRT; 160 AA.
AC 09K510;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Pertactin (P.68) (Fragment).
GN PRN.
OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
OC Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;
OC Bordetella.
OX NCBI_TaxID=518;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-9.73H+;
RX MEDLINE-20359389; PubMed-10899896;
RA Boursaux-Eude C., Guiso N.;
RT "Polymorphism of Repeated Regions of Pertactin in Bordetella
pertussis, Bordetella parapertussis, and Bordetella bronchiseptica.";
RL Infect. Immun. 68:4815-4817(2000).
DR EMBL; AJ250076; CAB76430.1; -
FT NON_TER 1
FT NON_TER 160
SQ SEQUENCE 160 AA; 15280 MW; FE9618ECD81013EB CRC64;

Query Match 94.9%; Score 232.5; DB 2; Length 160;
Best Local Similarity 90.2%; Pred. No. 6.1e-19;
Matches 46; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

OY 1 ORATIRRGDAPAGGAVPGGAVPGFGLDGMVGVDSSTVDLAQ 46
Db 54 ORATIRRGDAPAGGAVPGGAVPGFGLDGMVGVDSSTVDLAQ 104

RESULT 14
ID 09AHP1 PRELIMINARY; PRT; 179 AA.
AC 09AHP1;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Pertactin (Fragment).
GN PRN.
OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
OC Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;
OC Bordetella.
OX NCBI_TaxID=518;
RN [1]
RP SEQUENCE FROM N.A.
```

```

RC STRAIN-MBORD831;
RX MEDLINE-21117018; PubMed-11179374;
RA Register K.B.;
RT "Novel Genetic and Phenotypic Heterogeneity in Bordetella
   bronchiseptica pertactin.";
RL Infect. Immun. 69:1917-1921(2001).
DR EMBL: AF298589; AAK16690.1; -.
DR InterPro: IPR003992; pertactin.
DR PRINTS: PR01482; PERTACTIN.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 179 AA; 17270 MW; 325502FFB5483DF7 CRC64;

Query Match          94.9%; Score 232.5; DB 2; Length 179;
Best Local Similarity 90.2%; Pred. No. 6.9e-19;
Matches 46; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

QY 1 QRATIRRGDAPA----GGAVPGGAVPGGFGPLLDGWTGVDSSTVDLAQ 46
Db 79 QRATIRRGDAPAGAVPGGAVPGGFGPLLDGWTGVDSSTVDLAQ 129

RESULT 15
Q9KJY3 PRELIMINARY; PRT; 211 AA.
ID Q9KJY3
AC Q9KJY3;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE Pertactin (Fragment).
OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
OC Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;
OX NCBI_TaxID-518;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-B;
RA Kell D.J.; Fenwick B.;
RT "Variability in the pertactin genes of Bordetella bronchiseptica
   isolates includes regions coding for the CGXP and PGP families of
   amino acid repeats.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases;
DR EMBL: AF156769; AAF82393.1; -.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 211 AA; 20946 MW; E751PFA1510D99A4 CRC64;

Query Match          94.9%; Score 232.5; DB 2; Length 211;
Best Local Similarity 90.2%; Pred. No. 8.3e-19;
Matches 46; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

QY 1 QRATIRRGDAPA----GGAVPGGAVPGGFGPLLDGWTGVDSSTVDLAQ 46
Db 154 QRATIRRGDAPAGAVPGGAVPGGFGPLLDGWTGVDSSTVDLAQ 204

```

Search completed: March 17, 2003, 11:39:10
 Job time : 49.902 secs

GenCore version 5.1.4.P5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 17, 2003, 11:37:13 ; Search time 276 Seconds
(Without alignments)
107.456 Million cell updates/sec

Title: US-09-855-754B-8

Sequence: 1 GRATIRGDAPAGGAVPGCA.....PLDDGMYGVDSIVDLAQ 46

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

Pending_Patents_AA_Main:*

1: /cgn2_6/ptodata/2/paa/PCVUS_COMB.pep:*

2: /cgn2_6/ptodata/2/paa/US06_COMB.pep:*

3: /cgn2_6/ptodata/2/paa/US07_COMB.pep:*

4: /cgn2_6/ptodata/2/paa/US08_COMB.pep:*

5: /cgn2_6/ptodata/2/paa/US081_COMB.pep:*

6: /cgn2_6/ptodata/2/paa/US082_COMB.pep:*

7: /cgn2_6/ptodata/2/paa/US083_COMB.pep:*

8: /cgn2_6/ptodata/2/paa/US084_COMB.pep:*

9: /cgn2_6/ptodata/2/paa/US085_COMB.pep:*

10: /cgn2_6/ptodata/2/paa/US086_COMB.pep:*

11: /cgn2_6/ptodata/2/paa/US087_COMB.pep:*

12: /cgn2_6/ptodata/2/paa/US088_COMB.pep:*

13: /cgn2_6/ptodata/2/paa/US089_COMB.pep:*

14: /cgn2_6/ptodata/2/paa/US090_COMB.pep:*

15: /cgn2_6/ptodata/2/paa/US091_COMB.pep:*

16: /cgn2_6/ptodata/2/paa/US092_COMB.pep:*

17: /cgn2_6/ptodata/2/paa/US093_COMB.pep:*

18: /cgn2_6/ptodata/2/paa/US094_COMB.pep:*

19: /cgn2_6/ptodata/2/paa/US095_COMB.pep:*

20: /cgn2_6/ptodata/2/paa/US096_COMB.pep:*

21: /cgn2_6/ptodata/2/paa/US097_COMB.pep:*

22: /cgn2_6/ptodata/2/paa/US098_COMB.pep:*

23: /cgn2_6/ptodata/2/paa/US099_COMB.pep:*

24: /cgn2_6/ptodata/2/paa/US100_COMB.pep:*

25: /cgn2_6/ptodata/2/paa/US101_COMB.pep:*

26: /cgn2_6/ptodata/2/paa/US102_COMB.pep:*

27: /cgn2_6/ptodata/2/paa/US60_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	245	100.0	46	22	US-09-855-754-8
2	245	100.0	147	21	US-09-791-537-4547
3	245	100.0	150	21	US-09-791-537-4540
4	245	100.0	21	21	US-09-791-537-4562
5	245	100.0	154	21	US-09-791-537-4573
6	245	100.0	155	21	US-09-791-537-4568

7	245	100.0	911	21	US-09-791-537-37413	Sequence 37413, A
8	245	100.0	911	21	US-09-791-537-85698	Sequence 85698, A
9	245	100.0	911	22	US-09-855-754-4	Sequence 4, App11
10	232.5	94.9	51	22	US-09-855-754-7	Sequence 7, App11
11	232.5	94.9	158	21	US-09-791-537-4609	Sequence 4609, Ap
12	232.5	94.9	159	21	US-09-791-537-4603	Sequence 4603, Ap
13	232.5	94.9	160	21	US-09-791-537-4519	Sequence 4519, Ap
14	232.5	94.9	922	21	US-09-791-537-50286	Sequence 50286, A
15	232.5	94.9	922	22	US-09-855-754-6	Sequence 6, App11
16	219	89.4	910	22	US-09-855-754-13	Sequence 13, App1
17	214.5	87.6	51	22	US-09-791-537-75841	Sequence 75841, A
18	214.5	87.6	905	21	US-09-791-537-76981	Sequence 76981, A
19	214.5	87.6	905	21	US-09-791-537-76988	Sequence 76988, A
20	212	86.5	56	22	US-09-855-754-9	Sequence 9, App11
21	212	86.5	56	22	US-09-855-754-10	Sequence 10, App1
22	212	86.5	56	22	US-09-855-754-12	Sequence 12, App1
23	212	86.5	160	21	US-09-791-537-4640	Sequence 4640, Ap
24	212	86.5	161	21	US-09-791-537-4644	Sequence 4644, Ap
25	212	86.5	387	21	US-09-791-537-78243	Sequence 78243, A
26	212	86.5	910	21	US-09-791-537-21274	Sequence 21274, A
27	212	86.5	910	21	US-09-791-537-33443	Sequence 33443, A
28	212	86.5	910	21	US-09-791-537-112007	Sequence 112007, A
29	209.5	85.5	61	22	US-09-855-754-11	Sequence 11, App1
30	209.5	85.5	915	21	US-09-791-537-111629	Sequence 111629, A
31	208	84.9	907	21	US-09-791-537-78242	Sequence 78242, A
32	202	82.4	539	21	US-09-791-537-31601	Sequence 31601, A
33	202	82.4	910	21	US-09-791-537-98868	Sequence 98868, A
34	201.5	82.2	905	21	US-09-791-537-76981	Sequence 76981, A
35	201.5	82.2	745	14	US-09-010-928A-28	Sequence 28, App1
36	74.5	30.4	870	14	US-09-010-928A-2	Sequence 2, App1
37	73.5	30.0	350	19	US-09-595-328C-2166	Sequence 2166, Ap
38	73.5	30.0	438	19	US-09-513-996A-18485	Sequence 18485, A
39	73.5	30.0	438	20	US-09-620-394B-6666	Sequence 6666, Ap
40	73.5	30.0	439	19	US-09-535-328C-2185	Sequence 2185, Ap
41	73.5	30.0	460	19	US-09-513-996A-18484	Sequence 18484, A
42	73.5	30.0	460	20	US-09-620-394B-6665	Sequence 6665, Ap
43	73.5	30.0	461	19	US-09-595-328C-2164	Sequence 2164, Ap
44	73.5	30.0	646	19	US-09-513-996A-18483	Sequence 18483, A
45	73.5	30.0	646	20	US-09-620-394B-6664	Sequence 6664, Ap

ALIGNMENTS

RESULT 1
US-09-855-754-8
Sequence 8, Application US/09855754
GENERAL INFORMATION:
APPLICANT: BOURSAX-EDUE, CAROLINE
APPLICANT: GUISO-MACLOUF, NICOLE
TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED
TITLE OF INVENTION: REGIONS OF PERTACTIN IN BORDETELLA PERTUSSIS.
TITLE OF INVENTION: BORDETELLA PARAPERTUSSIS, AND BORDETELLA
TITLE OF INVENTION: BRONCHISEPTICA, THEIR USE IN DIAGNOSTICS, AND IN
FILE REFERENCE: 03495-0206-00000
CURRENT APPLICATION NUMBER: US/09/855,754
CURRENT FILING DATE: 2001-09-10
PRIOR APPLICATION NUMBER: 60/206,969
PRIOR FILING DATE: 2000-05-25
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 8
LENGTH: 46
TYPE: PRT
ORGANISM: Bordetella bronchiseptica
US-09-855-754-8

Query Match 100.0%; Score 245; DB 22; Length 46;
Best Local Similarity 100.0%; Pred. No. 3.2e-21;
Matches 46; Conservative 0; Mismatches 0; Indels 0;
Gy 1 GRATIRGDAPAGGAVPGGFPLLDDGMYGVDSIVDLAQ 46

DB 1 ORATIRGDAPAGAVPGGFGPGLDGMVGVDSSTVDLAQ 46
RESULT 2
US-09-791-537-4547
; Sequence 4547, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4547
; LENGTH: 147
; TYPE: PRF
; ORGANISM: Bordetella bronchiseptica
US-09-791-537-4547

Query Match 100.0%; Score 245; DB 21; Length 147;
Best Local Similarity 100.0%; Pred. No. 1.3e-20;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ORATIRGDAPAGAVPGGFGPGLDGMVGVDSSTVDLAQ 46
DB 47 ORATIRGDAPAGAVPGGFGPGLDGMVGVDSSTVDLAQ 92

RESULT 3
US-09-791-537-4540
; Sequence 4540, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4540
; LENGTH: 150
; TYPE: PRF
; ORGANISM: Bordetella bronchiseptica
US-09-791-537-4540

Query Match 100.0%; Score 245; DB 21; Length 150;
Best Local Similarity 100.0%; Pred. No. 1.3e-20;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ORATIRGDAPAGAVPGGFGPGLDGMVGVDSSTVDLAQ 46
DB 53 ORATIRGDAPAGAVPGGFGPGLDGMVGVDSSTVDLAQ 98

RESULT 4
US-09-791-537-4562
; Sequence 4562, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537

CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4562
; LENGTH: 150
; TYPE: PRF
; ORGANISM: Bordetella bronchiseptica
US-09-791-537-4562

Query Match 100.0%; Score 245; DB 21; Length 150;
Best Local Similarity 100.0%; Pred. No. 1.3e-20;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ORATIRGDAPAGAVPGGFGPGLDGMVGVDSSTVDLAQ 46
DB 50 ORATIRGDAPAGAVPGGFGPGLDGMVGVDSSTVDLAQ 95

RESULT 5
US-09-791-537-4573
; Sequence 4573, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4573
; LENGTH: 154
; TYPE: PRF
; ORGANISM: Bordetella bronchiseptica
US-09-791-537-4573

Query Match 100.0%; Score 245; DB 21; Length 154;
Best Local Similarity 100.0%; Pred. No. 1.3e-20;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ORATIRGDAPAGAVPGGFGPGLDGMVGVDSSTVDLAQ 46
DB 53 ORATIRGDAPAGAVPGGFGPGLDGMVGVDSSTVDLAQ 98

RESULT 6
US-09-791-537-4568
; Sequence 4568, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4568
; LENGTH: 155
; TYPE: PRF
; ORGANISM: Bordetella bronchiseptica
US-09-791-537-4568

Query Match 100.0%; Score 245; DB 21; Length 155;
Best Local Similarity 100.0%; Pred. No. 1.3e-20;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ORATIRGDAPAGAVPGGFGPGLDGMVGVDSSTVDLAQ 46


```
Db 54 ORATIRGDAPAGAVPGGFGFPLLDGMYGVDSSTVDLAQ 99
RESULT 7
US-09-791-537-37413
; Sequence 37413, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791.537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 37413
; LENGTH: 911
; TYPE: PRT
; ORGANISM: Bordetella bronchiseptica
US-09-791-537-37413
Query Match 100.0%; Score 245; DB 21; Length 911;
Best Local Similarity 100.0%; Pred. No. 1e-19;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 ORATIRGDAPAGAVPGGFGFPLLDGMYGVDSSTVDLAQ 46
Db 254 ORATIRGDAPAGAVPGGFGFPLLDGMYGVDSSTVDLAQ 299
RESULT 8
US-09-791-537-85698
; Sequence 85698, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791.537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 85698
; LENGTH: 911
; TYPE: PRT
; ORGANISM: Bordetella bronchiseptica
US-09-791-537-85698
Query Match 100.0%; Score 245; DB 21; Length 911;
Best Local Similarity 100.0%; Pred. No. 1e-19;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 ORATIRGDAPAGAVPGGFGFPLLDGMYGVDSSTVDLAQ 46
Db 254 ORATIRGDAPAGAVPGGFGFPLLDGMYGVDSSTVDLAQ 299
RESULT 9
US-09-855-754-4
; Sequence 4, Application US/09855754
; GENERAL INFORMATION:
; APPLICANT: BOURSAX-ENDE, CAROLINE
; APPLICANT: GUISSO-MACLOUF, NICOLE
; TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED
; TITLE OF INVENTION: REGIONS OF PERTACTIN IN BORDETELLA PERTUSSIS,
; TITLE OF INVENTION: BORDETELLA PARAPERTUSSIS, AND BORDETELLA
; TITLE OF INVENTION: BRONCHISEPTICA, THEIR USE IN DIAGNOSTICS, AND IN
; FILE REFERENCE: 03495-0206-00000
; ORGANISM: Bordetella bronchiseptica
US-09-855-754-4
Query Match 100.0%; Score 245; DB 22; Length 911;
Best Local Similarity 100.0%; Pred. No. 1e-19;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 ORATIRGDAPAGAVPGGFGFPLLDGMYGVDSSTVDLAQ 46
Db 254 ORATIRGDAPAGAVPGGFGFPLLDGMYGVDSSTVDLAQ 299
RESULT 10
US-09-855-754-7
; Sequence 7, Application US/09855754
; GENERAL INFORMATION:
; APPLICANT: BOURSAX-ENDE, CAROLINE
; APPLICANT: GUISSO-MACLOUF, NICOLE
; TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED
; TITLE OF INVENTION: REGIONS OF PERTACTIN IN BORDETELLA PERTUSSIS,
; TITLE OF INVENTION: BORDETELLA PARAPERTUSSIS, AND BORDETELLA
; TITLE OF INVENTION: BRONCHISEPTICA, THEIR USE IN DIAGNOSTICS, AND IN
; FILE REFERENCE: 03495-0206-00000
; CURRENT APPLICATION NUMBER: US/09/855.754
; CURRENT FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/206,969
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 7
; LENGTH: 51
; TYPE: PRT
; ORGANISM: Bordetella bronchiseptica
US-09-855-754-7
Query Match 94.9%; Score 232.5; DB 22; Length 51;
Best Local Similarity 90.2%; Pred. No. 1.1e-19;
Matches 46; Conservative 0; Mismatches 0; Indels 5; Gaps 1;
OY 1 ORATIRGDAPAGAVPGGFGFPLLDGMYGVDSSTVDLAQ 46
Db 254 ORATIRGDAPAGAVPGGFGFPLLDGMYGVDSSTVDLAQ 51
RESULT 11
US-09-791-537-4609
; Sequence 4609, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791.537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 4609
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Bordetella bronchiseptica
US-09-791-537-4609
```

Query Match 94.9%; Score 232.5; DB 21; Length 158;
Best Local Similarity 90.2%; Pred. No. 4e-19;
Matches 46; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

QY 1 QRATIRGDAPA-----GGAVPGAVPGGFGPILDDGMYGVDSSTVDLAQ 46
|||||
DB 54 QRATIRGDAPAGAVPGAVPGGFGPILDDGMYGVDSSTVDLAQ 104

RESULT 12
US-09-791-537-4603
; Sequence 4603, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4603
; LENGTH: 159
; TYPE: PRT
; ORGANISM: Bordetella bronchiseptica
US-09-791-537-4603

Query Match 94.9%; Score 232.5; DB 21; Length 159;
Best Local Similarity 90.2%; Pred. No. 4.1e-19;
Matches 46; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

QY 1 QRATIRGDAPA-----GGAVPGAVPGGFGPILDDGMYGVDSSTVDLAQ 46
|||||
DB 53 QRATIRGDAPAGAVPGAVPGGFGPILDDGMYGVDSSTVDLAQ 103

RESULT 13
US-09-791-537-4519
; Sequence 4519, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4519
; LENGTH: 160
; TYPE: PRT
; ORGANISM: Bordetella bronchiseptica
US-09-791-537-4519

Query Match 94.9%; Score 232.5; DB 21; Length 160;
Best Local Similarity 90.2%; Pred. No. 4.1e-19;
Matches 46; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

QY 1 QRATIRGDAPA-----GGAVPGAVPGGFGPILDDGMYGVDSSTVDLAQ 46
|||||
DB 54 QRATIRGDAPAGAVPGAVPGGFGPILDDGMYGVDSSTVDLAQ 104

RESULT 14
US-09-791-537-50286
; Sequence 50286, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.

APPLICANT: Debe, Derek
APPLICANT: Danzer, Joseph
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
TITLE OF INVENTION: METHODS OF USE THEREOF
FILE REFERENCE: 261/210
CURRENT APPLICATION NUMBER: US/09/791,537
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: PatentIn version 3.0
SEQ ID NO 50286
LENGTH: 922
TYPE: PRT
ORGANISM: Bordetella parapertussis
US-09-791-537-50286

Query Match 94.9%; Score 232.5; DB 21; Length 922;
Best Local Similarity 90.2%; Pred. No. 3.1e-18;
Matches 46; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

QY 1 QRATIRGDAPA-----GGAVPGAVPGGFGPILDDGMYGVDSSTVDLAQ 46
|||||
DB 254 QRATIRGDAPAGAVPGAVPGGFGPILDDGMYGVDSSTVDLAQ 304

RESULT 15
US-09-855-754-6
; Sequence 6, Application US/09855754
; GENERAL INFORMATION:
; APPLICANT: BOIRSAUX-EBDE, CAROLINE
; APPLICANT: GUISSO-MACLOFF, NICOLE
; TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED
; TITLE OF INVENTION: REGIONS OF PERTACTIN IN BORDETELLA PERTUSSIS,
; TITLE OF INVENTION: BORDETELLA PARAPERTUSSIS, AND BORDETELLA
; TITLE OF INVENTION: BRONCHISEPTICA, THEIR USE IN DIAGNOSTICS, AND IN
; FILE REFERENCE: 03495-0206-00000
; CURRENT APPLICATION NUMBER: US/09/855,754
; CURRENT FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/206,969
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 922
; TYPE: PRT
; ORGANISM: Bordetella parapertussis
US-09-855-754-6

Query Match 94.9%; Score 232.5; DB 22; Length 922;
Best Local Similarity 90.2%; Pred. No. 3.1e-18;
Matches 46; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

QY 1 QRATIRGDAPA-----GGAVPGAVPGGFGPILDDGMYGVDSSTVDLAQ 46
|||||
DB 254 QRATIRGDAPAGAVPGAVPGGFGPILDDGMYGVDSSTVDLAQ 304

Search completed: March 17, 2003, 11:44:55
Job time : 277 secs

GenCore version 5.1.4.P5-4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 17, 2003, 11:37:38 ; Search time 46.902 Seconds
(without alignments)
127.976 Million cell updates/sec

Title: US-09-855-754B-8
Perfect score: 245
Sequence: 1 ORATIRGDAPAGGAVPGGA.....PLDGYGVDSYDVLDAQ 46

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 603988 seqs, 130485580 residues

Total number of hits satisfying chosen parameters: 603988

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_AA_New:*
1: /cgn2-6/ptodata/1/paa/PCF_NEW.COM3.pep:*
2: /cgn2-6/ptodata/1/paa/US06_NEW.COM3.pep:*
3: /cgn2-6/ptodata/1/paa/US07_NEW.COM3.pep:*
4: /cgn2-6/ptodata/1/paa/US08_NEW.COM3.pep:*
5: /cgn2-6/ptodata/1/paa/US09_NEW.COM3.pep:*
6: /cgn2-6/ptodata/1/paa/US10_NEW.COM3.pep:*
7: /cgn2-6/ptodata/1/paa/US60_NEW.COM3.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	245	100.0	46 5 US-09-855-754B-8	Sequence 8, Appli
2	245	100.0	911 5 US-09-855-754B-4	Sequence 4, Appli
3	245	100.0	911 6 US-10-227-353-4	Sequence 4, Appli
4	232.5	94.9	51 5 US-09-855-754B-7	Sequence 7, Appli
5	232.5	94.9	922 6 US-09-855-754B-6	Sequence 6, Appli
6	232.5	94.9	922 6 US-10-227-353-6	Sequence 5, Appli
7	219	89.4	910 5 US-09-855-754B-5	Sequence 5, Appli
8	214.5	87.6	51 5 US-09-855-754B-13	Sequence 13, Appli
9	212	86.5	56 5 US-09-855-754B-9	Sequence 9, Appli
10	212	86.5	56 5 US-09-855-754B-10	Sequence 10, Appli
11	212	86.5	56 5 US-09-855-754B-12	Sequence 12, Appli
12	212	86.5	768 6 US-10-282-122A-51012	Sequence 51012, A
13	209.5	85.5	61 5 US-09-855-754B-11	Sequence 11, Appli
14	202	82.4	910 6 US-10-227-353-2	Sequence 2, Appli
15	202	82.4	910 6 US-10-366-683-26256	Sequence 26256, A
16	76	29.0	342 6 US-10-366-683-26256	Sequence 26256, A
17	66	26.9	171 1 PCT-US02-18256-45	Sequence 45, Appli
18	66	26.9	3420 1 PCT-US02-18256-51	Sequence 51, Appli
19	65	26.5	177 6 US-10-366-683-27635	Sequence 27635, A
20	64.5	26.3	412 6 US-10-224-999A-3466	Sequence 3466, Ap
21	64	26.1	278 6 US-10-282-122A-49688	Sequence 49688, A
22	62	25.3	63 1 PCT-US02-18256-46	Sequence 46, Appli
23	62	25.3	315 6 US-10-366-683-31381	Sequence 31381, A
24	62	25.3	2016 1 PCT-US02-18256-52	Sequence 52, Appli
25	61	24.9	229 6 US-10-366-683-27802	Sequence 27802, A
26	60.5	24.7	461 6 US-10-366-683-24717	Sequence 24717, A
			310 6 US-10-366-683-16897	Sequence 16897, A

27	60	24.5	47	6	US-10-276-774-2507	Sequence 2507, Ap
28	60	24.5	217	6	US-10-366-683-24723	Sequence 24723, A
29	59.5	24.3	264	6	US-10-366-683-16625	Sequence 16625, A
30	59.5	24.3	399	6	US-10-310-154-508	Sequence 508, App
31	59	24.1	130	6	US-10-366-683-23606	Sequence 23606, A
32	59	24.1	162	1	PCT-US02-32727-24021	Sequence 24021, A
33	59	24.1	162	5	US-09-978-825-24021	Sequence 24021, A
34	59	24.1	162	6	US-10-057-458-24021	Sequence 24021, A
35	59	24.1	426	1	PCT-US02-32727-4545	Sequence 4545, Ap
36	59	24.1	426	5	US-09-978-825-4545	Sequence 4545, Ap
37	59	24.1	426	6	US-10-057-458-4545	Sequence 4545, Ap
38	59	24.1	714	1	PCT-US02-32727-30064	Sequence 30064, A
39	59	24.1	714	5	US-09-978-825-30064	Sequence 30064, A
40	59	24.1	792	6	US-10-366-683-25823	Sequence 25823, A
41	58.5	23.9	460	6	US-10-282-122A-62312	Sequence 62312, A
42	58.5	23.9	480	6	US-10-282-122A-64902	Sequence 64902, A
43	58.5	23.9	1175	6	US-10-366-683-25044	Sequence 25044, A
44	58	23.7	149	6	US-10-366-683-22850	Sequence 22850, A
45	58	23.7	176	6	US-10-282-122A-43022	Sequence 43022, A

ALIGNMENTS

```

RESULT 1
US-09-855-754B-8
Sequence 8, Application US/09855754B
GENERAL INFORMATION:
APPLICANT: BOURSAX-ETUDE, CAROLINE
APPLICANT: GUISO-MACLOUF, NICOLE
TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED
TITLE OF INVENTION: REGIONS OF PERACTIN IN BORDETELLA PERTUSSIS,
TITLE OF INVENTION: BORDETELLA PARAPERTUSSIS, AND BORDETELLA
TITLE OF INVENTION: BRONCHISEPTICA, THEIR USE IN DIAGNOSTICS, AND IN
FILE REFERENCE: 03495-0206-00000
CURRENT APPLICATION NUMBER: US/09/855,754B
PRIOR FILING DATE: 2001-05-16
PRIOR APPLICATION NUMBER: 60/206,969
NUMBER OF SEQ ID NOS: 25
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 8
LENGTH: 46
TYPE: PRT
ORGANISM: Bordetella bronchiseptica
US-09-855-754B-8
Query Match 100.0%; Score 245; DB 5; Length 46;
Best Local Similarity 100.0%; Pred. No. 2e-19;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ORATIRGDAPAGGAVPGGAPGRLDGYGVDSYDVLDAQ 46
DB 1 ORATIRGDAPAGGAVPGGAPGRLDGYGVDSYDVLDAQ 46
RESULT 2
US-09-855-754B-4
Sequence 4, Application US/09855754B
GENERAL INFORMATION:
APPLICANT: BOURSAX-ETUDE, CAROLINE
APPLICANT: GUISO-MACLOUF, NICOLE
TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED
TITLE OF INVENTION: REGIONS OF PERACTIN IN BORDETELLA PERTUSSIS,
TITLE OF INVENTION: BORDETELLA PARAPERTUSSIS, AND BORDETELLA
TITLE OF INVENTION: BRONCHISEPTICA, THEIR USE IN DIAGNOSTICS, AND IN
FILE REFERENCE: 03495-0206-00000
CURRENT APPLICATION NUMBER: US/09/855,754B
PRIOR FILING DATE: 2001-05-16
PRIOR APPLICATION NUMBER: 60/206,969
PRIOR FILING DATE: 2000-05-25

```

NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 4
; LENGTH: 911
; TYPE: PRT
; ORGANISM: Bordetella bronchiseptica
US-09-855-754B-4

Query Match 100.0%; Score 245; DB 5; Length 911;
Best Local Similarity 100.0%; Pred. No. 4.3e-18;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ORATIRRGDAPAGGAVPGGFGFPLDGMVGVDSSTVDLAQ 46
DB 254 ORATIRRGDAPAGGAVPGGFGFPLDGMVGVDSSTVDLAQ 299

RESULT 3

US-10-227-353-4

; Sequence 4, Application US/10227353

; GENERAL INFORMATION:

; APPLICANT: CLARE, JEFFREY J.

; ROMANOS, MICHAEL A.

; TITLE OF INVENTION: EXPRESSION OF HETEROLOGOUS PROTEIN IN YEAST

; NUMBER OF SEQUENCES: 17

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Millen, White, Zelano & Branigan, P.C.

; STREET: 2200 Clarendon Blvd., Suite 1400

; CITY: ARLINGTON

; STATE: VA

; COUNTRY: USA

; ZIP: 22201

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentln Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/227,353

; FILING DATE: 26-Aug-2002

; PRIORITY FILING DATE:

; APPLICATION NUMBER: US/08/460,269C

; FILING DATE: 02-Jun-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Lebovitz, Richard M.

; REGISTRATION NUMBER: 37,067

; REFERENCE/DOCKET NUMBER: Popov-2

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (703) 243-6333

; TELEFAX: (703) 243-6410

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 911 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; SEQUENCE DESCRIPTION: SEQ ID NO: 4:

; US-10-227-353-4

; Query Match 100.0%; Score 245; DB 6; Length 911;

; Best Local Similarity 100.0%; Pred. No. 4.3e-18;

; Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ORATIRRGDAPAGGAVPGGFGFPLDGMVGVDSSTVDLAQ 46
DB 254 ORATIRRGDAPAGGAVPGGFGFPLDGMVGVDSSTVDLAQ 299

RESULT 4
US-09-855-754B-7
; Sequence 7, Application US/09855754B
; GENERAL INFORMATION:

APPLICANT: BOURSAX-EUDE, CAROLINE
; APPLICANT: GUISSO-MACLOUE, NICOLE
; TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED
; TITLE OF INVENTION: REGIONS OF PERTACTIN IN BORDETELLA PERTUSSIS,
; TITLE OF INVENTION: BORDETELLA PARAPERTUSSIS, AND BORDETELLA
; TITLE OF INVENTION: BRONCHISEPTICA, THEIR USE IN DIAGNOSTICS, AND IN
; FILE REFERENCE: 03495-0206-00000
; CURRENT APPLICATION NUMBER: US/09/855,754B
; CURRENT FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 60/206,969
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 7
; LENGTH: 51
; TYPE: PRT
; ORGANISM: Bordetella bronchiseptica
US-09-855-754B-7

Query Match 94.9%; Score 232.5; DB 5; Length 51;
Best Local Similarity 90.2%; Pred. No. 4.9e-18;
Matches 46; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

QY 1 ORATIRRGDAPAGGAVPGGFGFPLDGMVGVDSSTVDLAQ 46
DB 254 ORATIRRGDAPAGGAVPGGFGFPLDGMVGVDSSTVDLAQ 51

RESULT 5

US-09-855-754B-6

; Sequence 6, Application US/09855754B

; GENERAL INFORMATION:

; APPLICANT: BOURSAX-EUDE, CAROLINE

; APPLICANT: GUISSO-MACLOUE, NICOLE

; TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED

; TITLE OF INVENTION: REGIONS OF PERTACTIN IN BORDETELLA PERTUSSIS,
; TITLE OF INVENTION: BORDETELLA PARAPERTUSSIS, AND BORDETELLA
; TITLE OF INVENTION: BRONCHISEPTICA, THEIR USE IN DIAGNOSTICS, AND IN
; FILE REFERENCE: 03495-0206-00000
; CURRENT APPLICATION NUMBER: US/09/855,754B
; CURRENT FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 60/206,969
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 6
; LENGTH: 922
; TYPE: PRT
; ORGANISM: Bordetella parapertussis
US-09-855-754B-6

Query Match 94.9%; Score 232.5; DB 5; Length 922;
Best Local Similarity 90.2%; Pred. No. 9.4e-17;
Matches 46; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

QY 1 ORATIRRGDAPAGGAVPGGFGFPLDGMVGVDSSTVDLAQ 46
DB 254 ORATIRRGDAPAGGAVPGGFGFPLDGMVGVDSSTVDLAQ 304

RESULT 6
US-10-227-353-6
; Sequence 6, Application US/10227353
; GENERAL INFORMATION:

; APPLICANT: CLARE, JEFFREY J.

; ROMANOS, MICHAEL A.

; TITLE OF INVENTION: EXPRESSION OF HETEROLOGOUS PROTEIN IN YEAST

; NUMBER OF SEQUENCES: 17

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Millen, White, Zelano & Branigan, P.C.

STREET: 2200 Clarendon Blvd., Suite 1400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22201
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/227,353
FILING DATE: 26-Aug-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/460,269C
FILING DATE: 02-Jun-1995
ATTORNEY/AGENT INFORMATION:
NAME: Lebovitz, Richard M.
REGISTRATION NUMBER: 37,067
REFERENCE/DOCKET NUMBER: Popov-2
TELEPHONE: (703) 243-6333
TELEFAX: (703) 243-6410
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 922 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-227-353-6

Query Match 94.9%; Score 232.5; DB 6; Length 922;
Best Local Similarity 90.2%; Pred. No. 9.4e-17;
Matches 46; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

OY 1 ORATIRRGDAPAGAVPGAVGAVGCGFPLLDGWTGVVSDSTVDLAQ 46
Db 254 ORATIRRGDAPAGAVPGAVGAVGCGFPLLDGWTGVVSDSTVDLAQ 304

RESULT 7
US-09-855-754B-5
Sequence 5, Application US/09855754B
GENERAL INFORMATION:
APPLICANT: BOURSAX-EBUE, CAROLINE
APPLICANT: GUIISO-MACLOUF, NICOLE
TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED
REGIONS OF PERTACTIN IN BORDETELLA PERTUSSIS,
TITLE OF INVENTION: BORDETELLA PARAPERTUSSIS, AND BORDETELLA
TITLE OF INVENTION: BRONCHISEPTICA, THEIR USE IN DIAGNOSTICS, AND IN
TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS
FILE REFERENCE: 03495-0206-00000
CURRENT APPLICATION NUMBER: US/09/855,754B
CURRENT FILING DATE: 2001-05-16
PRIOR APPLICATION NUMBER: 60/206,969
PRIOR FILING DATE: 2000-05-25
NUMBER OF SEQ ID NOS: 25
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 5
LENGTH: 910
TYPE: PRT
ORGANISM: Bordetella pertussis
US-09-855-754B-5

Query Match 89.4%; Score 219; DB 5; Length 910;
Best Local Similarity 76.8%; Pred. No. 2.6e-15;
Matches 43; Conservative 3; Mismatches 0; Indels 10; Gaps 1;

OY 1 ORATIRRGDAPAGAVPGAVGAVGCGFPLLDGWTGVVSDSTVDLAQ 46
Db 254 ORATIRRGDAPAGAVPGAVGAVGCGFPLLDGWTGVVSDSTVDLAQ 309

RESULT 8
US-09-855-754B-13
Sequence 13, Application US/09855754B
GENERAL INFORMATION:
APPLICANT: BOURSAX-EBUE, CAROLINE
APPLICANT: GUIISO-MACLOUF, NICOLE
TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED
REGIONS OF PERTACTIN IN BORDETELLA PERTUSSIS,
TITLE OF INVENTION: BORDETELLA PARAPERTUSSIS, AND BORDETELLA
TITLE OF INVENTION: BRONCHISEPTICA, THEIR USE IN DIAGNOSTICS, AND IN
TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS
FILE REFERENCE: 03495-0206-00000
CURRENT APPLICATION NUMBER: US/09/855,754B
CURRENT FILING DATE: 2001-05-16
PRIOR APPLICATION NUMBER: 60/206,969
PRIOR FILING DATE: 2000-05-25
NUMBER OF SEQ ID NOS: 25
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 13
LENGTH: 51
TYPE: PRT
ORGANISM: Bordetella bronchiseptica
US-09-855-754B-13

Query Match 87.6%; Score 214.5; DB 5; Length 51;
Best Local Similarity 82.4%; Pred. No. 4.1e-16;
Matches 42; Conservative 3; Mismatches 1; Indels 5; Gaps 1;

OY 1 ORATIRRGDAPAGAVPGAVGAVGCGFPLLDGWTGVVSDSTVDLAQ 46
Db 1 ORATIRRGDAPAGAVPGAVGAVGCGFPLLDGWTGVVSDSTVDLAQ 51

RESULT 9
US-09-855-754B-9
Sequence 9, Application US/09855754B
GENERAL INFORMATION:
APPLICANT: BOURSAX-EBUE, CAROLINE
APPLICANT: GUIISO-MACLOUF, NICOLE
TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED
REGIONS OF PERTACTIN IN BORDETELLA PERTUSSIS,
TITLE OF INVENTION: BORDETELLA PARAPERTUSSIS, AND BORDETELLA
TITLE OF INVENTION: BRONCHISEPTICA, THEIR USE IN DIAGNOSTICS, AND IN
TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS
FILE REFERENCE: 03495-0206-00000
CURRENT APPLICATION NUMBER: US/09/855,754B
CURRENT FILING DATE: 2001-05-16
PRIOR APPLICATION NUMBER: 60/206,969
PRIOR FILING DATE: 2000-05-25
NUMBER OF SEQ ID NOS: 25
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 9
LENGTH: 56
TYPE: PRT
ORGANISM: Bordetella bronchiseptica
US-09-855-754B-9

Query Match 86.5%; Score 212; DB 5; Length 56;
Best Local Similarity 75.0%; Pred. No. 8.4e-16;
Matches 42; Conservative 2; Mismatches 2; Indels 10; Gaps 1;

OY 1 ORATIRRGDAPAGAVPGAVGAVGCGFPLLDGWTGVVSDSTVDLAQ 46
Db 1 ORATIRRGDAPAGAVPGAVGAVGCGFPLLDGWTGVVSDSTVDLAQ 56

RESULT 10
US-09-855-754B-10
Sequence 10, Application US/09855754B
GENERAL INFORMATION:
APPLICANT: BOURSAX-EBUE, CAROLINE
APPLICANT: GUIISO-MACLOUF, NICOLE

```

; TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED
; TITLE OF INVENTION: REGIONS OF PERTACTIN IN BORDETELLA PERTUSSIS,
; TITLE OF INVENTION: BORDETELLA PARAPERTUSSIS, AND BORDETELLA
; TITLE OF INVENTION: BRONCHISEPTICA, THEIR USE IN DIAGNOSTICS, AND IN
; FILE REFERENCE: 03495-0206-00000
; CURRENT APPLICATION NUMBER: US/09/855,754B
; PRIOR FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 60/206,969
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Bordetella bronchiseptica
US-09-855-754B-10

Query Match
Best Local Similarity 86.5%; Score 212; DB 5; Length 56;
Matches 42; Conservative 3; Mismatches 1; Indels 10; Gaps 1;

Oy 1 ORATIRRGDAPAGGAVPGGAV-----PGGFGPLDGYGVDSSTVDLAQ 46
Db 1 ORATIRRGDAPAGGAVPGGAVPGGFGPGGFGPVLDMGTGVDSGSSVELAQ 56

RESULT 11
US-09-855-754B-12
; Sequence 12, Application US/09855754B
; GENERAL INFORMATION:
; APPLICANT: BOURSAX-ENDE, CAROLINE
; APPLICANT: GUISO-MACLOUF, NICOLE
; TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED
; TITLE OF INVENTION: REGIONS OF PERTACTIN IN BORDETELLA PERTUSSIS,
; TITLE OF INVENTION: BORDETELLA PARAPERTUSSIS, AND BORDETELLA
; TITLE OF INVENTION: BRONCHISEPTICA, THEIR USE IN DIAGNOSTICS, AND IN
; FILE REFERENCE: 03495-0206-00000
; CURRENT APPLICATION NUMBER: US/09/855,754B
; PRIOR FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 60/206,969
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Bordetella bronchiseptica
US-09-855-754B-12

Query Match
Best Local Similarity 86.5%; Score 212; DB 5; Length 56;
Matches 42; Conservative 3; Mismatches 1; Indels 10; Gaps 1;

Oy 1 ORATIRRGDAPAGGAVPGGAV-----PGGFGPLDGYGVDSSTVDLAQ 46
Db 1 ORATIRRGDAPAGGAVPGGAVPGGFGPGGFGPVLDMGTGVDSGSSVELAQ 56

RESULT 12
US-10-282-122A-51012
; Sequence 51012, Application US/10282122A
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant

```

```

; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; Remaining prior Application data removed - See file wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 51012
; LENGTH: 768
; TYPE: PRT
; ORGANISM: Bordetella pertussis
US-10-282-122A-51012

Query Match
Best Local Similarity 86.5%; Score 212; DB 6; Length 768;
Matches 42; Conservative 3; Mismatches 1; Indels 10; Gaps 1;

Oy 1 ORATIRRGDAPAGGAVPGGAV-----PGGFGPLDGYGVDSSTVDLAQ 46
Db 112 ORATIRRGDAPAGGAVPGGAVPGGFGPGGFGPVLDMGTGVDSGSSVELAQ 167

RESULT 13
US-09-855-754B-11
; Sequence 11, Application US/09855754B
; GENERAL INFORMATION:
; APPLICANT: BOURSAX-ENDE, CAROLINE
; APPLICANT: GUISO-MACLOUF, NICOLE
; TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED
; TITLE OF INVENTION: REGIONS OF PERTACTIN IN BORDETELLA PERTUSSIS,
; TITLE OF INVENTION: BORDETELLA PARAPERTUSSIS, AND BORDETELLA
; TITLE OF INVENTION: BRONCHISEPTICA, THEIR USE IN DIAGNOSTICS, AND IN
; FILE REFERENCE: 03495-0206-00000
; CURRENT APPLICATION NUMBER: US/09/855,754B
; PRIOR FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 60/206,969
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 61
; TYPE: PRT
; ORGANISM: Bordetella bronchiseptica
US-09-855-754B-11

Query Match
Best Local Similarity 85.5%; Score 209.5; DB 5; Length 61;
Matches 42; Conservative 3; Mismatches 1; Indels 15; Gaps 1;

```

us-09-855-754b-8.ram

Page 5

CURRENT FILING DATE: 2003-02-13
PRIOR APPLICATION NUMBER: 09/252,991
PRIOR FILING DATE: 1999-02-18
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 26256
LENGTH: 542
TYPE: PRT

US-10-366-683-26256 ORGANISM: Pseudomonas aeruginosa
Query Match 29.0%; Score 100.00

Query Match	29.0%;	Score 71;	DB 6;	Length 542;
Best Local Similarity	51.6%;	Pred. No. 10;		
Matches 16; Conservative	0;	Mismatches 11;	Indels 4;	Gaps 1;

Matches 16; Conservative 0; Mismatches 11; Indels 4; Gaps 1;

	27	SCHOOL	6	HIGHWAYS	11	INDUST	4	GAPS	17
4	TIRRGDAPAGGAVPGGFPLLDGWY	34							

4 TIRRGDAPAGGAVPGGAVGGFGPLLDGWYG 34
| | | | |
1 TAGRGRDPAGGVVVGRRPPGG---AHGWRG 27

U

Search completed: March 17, 2003, 11:45:56

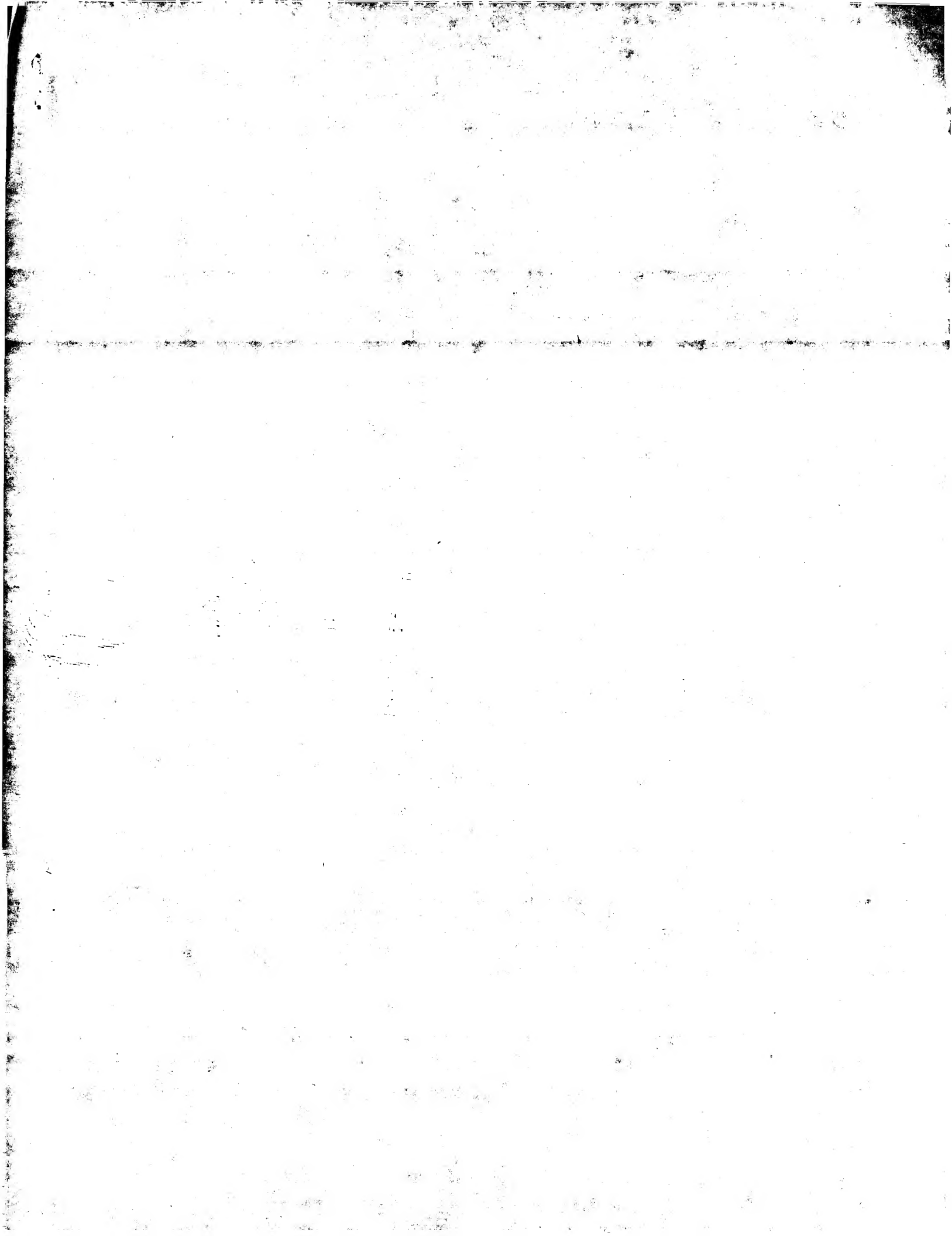
```

search completed: March 17, 2003, 11:45:56
job time : 47.902 secs

```

Job time : 47.902 secs

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84




```

Sequence 25 Application US/09855754B
GENERAL INFORMATION:
APPLICANT: BOURSAX-EBDE, CAROLINE
APPLICANT: GUISO-MACLOUF, NICOLE
TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED
TITLE OF INVENTION: REGIONS OF PERTACTIN IN BORDETELLA PERTUSSIS,
TITLE OF INVENTION: BORDETELLA PARAPERTUSSIS, AND BORDETELLA
TITLE OF INVENTION: BRONCHISEPTICA, THEIR USE IN DIAGNOSTICS, AND IN
TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS
FILE REFERENCE: 03495-0206-00000
CURRENT APPLICATION NUMBER: US/09/855.754B
CURRENT FILING DATE: 2001-05-16
PRIOR APPLICATION NUMBER: 60/206,969
PRIOR FILING DATE: 2000-05-25
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 25
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic peptide
FEATURE:
NAME/KEY: MOD_RES
LOCATION: (3)..(4)
OTHER INFORMATION: Phe Asp, Phe Gly or Ala Val
US-09-855-754B-25
GGXXPI

```


GenCore version 5.1.4-p5-4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 17, 2003, 12:05:19 ; Search time 31.3333 Seconds
(without alignments)
21.263 Million cell updates/sec

Title: SEQ25-AV

Perfect score: 27
Sequence: 1 ggapv 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A.Geneseq_101002.*
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	27	100.0	10	23	AAE17143
2	27	100.0	15	23	AAE17145
3	27	100.0	15	23	AAE17160
4	27	100.0	15	23	AAE17161
5	27	100.0	15	23	AAE17162
6	27	100.0	15	23	AAE17163
7	27	100.0	20	23	AAE17144
8	27	100.0	23	23	AAU89387
9	27	100.0	31	23	AAE17173
10	27	100.0	36	23	AAE17141

11	27	100.0	36	23	AAE17170
12	27	100.0	36	23	AAE17172
13	27	100.0	36	23	AAE17174
14	27	100.0	38	15	AAE56849
15	27	100.0	38	15	AAE56850
16	27	100.0	41	23	AAE17171
17	27	100.0	45	23	AAE17153
18	27	100.0	45	23	AAE17154
19	27	100.0	45	23	AAE17155
20	27	100.0	45	23	AAE17156
21	27	100.0	45	23	AAE17157
22	27	100.0	45	23	AAE17158
23	27	100.0	45	23	AAE17159
24	27	100.0	46	23	AAE16187
25	27	100.0	47	22	ABE15654
26	27	100.0	51	23	AAE16186
27	27	100.0	51	23	AAE16192
28	27	100.0	54	23	AAE17150
29	27	100.0	54	23	AAE17151
30	27	100.0	56	23	AAE16188
31	27	100.0	56	23	AAE16189
32	27	100.0	56	23	AAE16191
33	27	100.0	56	23	AAE17152
34	27	100.0	59	22	AAU55300
35	27	100.0	59	23	AAE17147
36	27	100.0	59	23	AAE17149
37	27	100.0	61	23	AAE16190
38	27	100.0	64	23	AAE17148
39	27	100.0	65	22	ABG27823
40	27	100.0	76	21	AAE61538
41	27	100.0	84	22	AAU58355
42	27	100.0	91	22	AAO03675
43	27	100.0	95	21	AAE61537
44	27	100.0	96	23	AAO21287
45	27	100.0	104	23	ABE06381

ALIGNMENTS

RESULT 1
ID AAE17143 standard; peptide; 10 AA.
XX AAE17143:
XX 18-APR-2002 (first entry)
XX
XX Bordetella pertussis pertactin (Prn) related peptide #2.
DE
DE Pertactin: prn; antibacterial; immunostimulant; antimicrobial; vaccine;
KW diphtheria; tetanus; polio; Haemophilus influenza b infection; therapy;
KW immune response.
XX
XX Unidentified.
OS
OS WO200200695-A2.
PN
PN 03-JAN-2002.
PD
PD 29-JUN-2001; 2001WO-NL00493.
PF
PF 30-JUN-2000; 2000EP-0202309.
PR
PR (NEMO-) NEDERLANDEN MIN WELZIJN.
PA
PA Mooi FR;
XX
XX WPI; 2002-139897/18.
XX
XX New polypeptides derived from Bordetella pertussis pertactin, useful as
PT a vaccine against infections caused by Bordetella strains, and other
PT infectious diseases of mammals, e.g. diphtheria, tetanus, or polio

XX PS Disclosure; Page 7; 52pp; English.
 CC CC The invention relates to polypeptides derived from Bordetella pertussis
 CC CC pertactin (Pn1). The polypeptide is useful in the preparation of
 CC CC vaccines against B. pertussis, B. parapertussis, B. bronchiseptica and
 CC CC other infectious diseases of mammals including diphtheria, tetanus, polio
 CC CC and infections caused by Haemophilus influenza b. The polypeptide is
 CC CC especially useful for eliciting an immune response against Bordetella sp.
 CC CC Antibodies against the polypeptide may be used for pharmaceutical and/or
 CC CC diagnostic purposes, particularly for treating or preventing infections
 CC CC caused by Bordetella pertussis or Bordetella parapertussis. The present
 CC CC sequence is B. pertussis pml related peptide used in the invention.
 XX SQ Sequence 10 AA;
 Query Match 100.0%; Score 27; DB 23; Length 10;
 Best Local Similarity 100.0%; Pred. No. 32;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 GGAVP 5
 DB 1 GGAVP 5
 RESULT 2
 AAE17145
 ID AAE17145 standard; peptide; 15 AA.
 XX AC AAE17145;
 XX DT 18-APR-2002 (first entry)
 XX DE Bordetella pertussis pertactin (Pn1) related peptide #4.
 XX KW Pertactin; pml; antibacterial; immunostimulant; antimicrobial; vaccine;
 KW diphtheria; tetanus; polio; Haemophilus influenza b infection; therapy;
 KW immune response.
 XX OS Unidentified.
 XX PN WO200200695-A2.
 XX PD 03-JAN-2002.
 XX PE 29-JUN-2001; 2001WO-NL00493.
 XX PR 30-JUN-2000; 2000EP-0202309.
 XX PA (NEME-) NEDERLANDEN MIN WELZIJN.
 XX PI Mool FR;
 XX PT WPI; 2002-139897/18.
 XX DR New polypeptides derived from Bordetella pertussis pertactin, useful as
 XX PT a vaccine against infections caused by Bordetella strains, and other
 XX PT infectious diseases of mammals, e.g. diphtheria, tetanus, or polio
 XX PS Claim 6; Page 7; 52pp; English.
 XX CC The invention relates to polypeptides derived from Bordetella pertussis
 CC CC pertactin (Pn1). The polypeptide is useful in the preparation of
 CC CC vaccines against B. pertussis, B. parapertussis, B. bronchiseptica and
 CC CC other infectious diseases of mammals including diphtheria, tetanus, polio
 CC CC and infections caused by Haemophilus influenza b. The polypeptide is
 CC CC especially useful for eliciting an immune response against Bordetella sp.
 CC CC Antibodies against the polypeptide may be used for pharmaceutical and/or
 CC CC diagnostic purposes, particularly for treating or preventing infections
 CC CC caused by Bordetella pertussis or Bordetella parapertussis. The present
 CC CC sequence is B. pertussis pml related peptide used in the invention.
 XX SQ Sequence 15 AA;

Query Match 100.0%; Score 27; DB 23; Length 15;
 Best Local Similarity 100.0%; Pred. No. 47;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 GGAVP 5
 DB 1 GGAVP 5
 RESULT 3
 AAE17160
 ID AAE17160 standard; peptide; 15 AA.
 XX AC AAE17160;
 XX DT 18-APR-2002 (first entry)
 XX DE Monoclonal Ab derived Pcp1 peptide specific to region 1 of pertactin.
 XX KW Pertactin; pml; antibacterial; immunostimulant; antimicrobial; vaccine;
 KW diphtheria; tetanus; polio; Haemophilus influenza b infection; therapy;
 KW immune response.
 XX OS Unidentified.
 XX PN WO200200695-A2.
 XX PD 03-JAN-2002.
 XX PE 29-JUN-2001; 2001WO-NL00493.
 XX PR 30-JUN-2000; 2000EP-0202309.
 XX PA (NEME-) NEDERLANDEN MIN WELZIJN.
 XX PI Mool FR;
 XX PT WPI; 2002-139897/18.
 XX DR New polypeptides derived from Bordetella pertussis pertactin, useful as
 XX PT a vaccine against infections caused by Bordetella strains, and other
 XX PT infectious diseases of mammals, e.g. diphtheria, tetanus, or polio
 XX PS Example 1; Fig 2b; 52pp; English.
 XX CC The invention relates to polypeptides derived from Bordetella pertussis
 CC CC pertactin (Pn1). The polypeptide is useful in the preparation of
 CC CC vaccines against B. pertussis, B. parapertussis, B. bronchiseptica and
 CC CC other infectious diseases of mammals including diphtheria, tetanus, polio
 CC CC and infections caused by Haemophilus influenza b. The polypeptide is
 CC CC especially useful for eliciting an immune response against Bordetella sp.
 CC CC Antibodies against the polypeptide may be used for pharmaceutical and/or
 CC CC diagnostic purposes, particularly for treating or preventing infections
 CC CC caused by Bordetella pertussis or Bordetella parapertussis. The present
 CC CC sequence is a peptide derived from monoclonal antibody which is specific
 CC CC to region 1 of pertactin.
 XX SQ Sequence 15 AA;
 Query Match 100.0%; Score 27; DB 23; Length 15;
 Best Local Similarity 100.0%; Pred. No. 47;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 GGAVP 5
 DB 11 GGAVP 15
 RESULT 4
 AAE17161
 ID AAE17161 standard; peptide; 15 AA.

AC	AAE17161;
XX	
DT	18-APR-2002 (first entry)
DE	Monoclonal Ab derived Pep2 peptide specific to region 1 of pertactin.
XX	
KW	Pertactin; prnl; antibacterial; immunostimulant; antimicrobial; vaccine; diphtheria; tetanus; polio; Haemophilus influenza b infection; therapy;
KW	immune response.
XX	
OS	Unidentified.
PM	MO200200695-A2.
PD	03-JAN-2002.
XX	
PF	29-JUN-2001; 2001WO-NL00493.
PR	30-JUN-2000; 2000EP-0202309.
XX	
PA	(NEMO-) NEDERLANDEN MIN WELZIJN.
PL	Mool FR;
DR	WPI; 2002-139897/18.
PT	New polypeptides derived from Bordetella pertussis pertactin, useful as a vaccine against infections caused by Bordetella strains, and other infectious diseases of mammals, e.g. diphtheria, tetanus, or polio - Example 1; Fig 2b; 52pp; English.
XX	
CC	The invention relates to polypeptides derived from Bordetella pertussis pertactin (Prnl). The polypeptide is useful in the preparation of vaccines against B. pertussis, B. paraptussis, B. bronchiseptica and other infectious diseases of mammals including diphtheria, tetanus, polio and infections caused by Haemophilus influenza b. The polypeptide is especially useful for eliciting an immune response against Bordetella sp. Antibodies against the polypeptide may be used for pharmaceutical and/or diagnostic purposes, particularly for treating or preventing infections caused by Bordetella pertussis or Bordetella parapertussis. The present sequence is a peptide derived from monoclonal antibody which is specific to region 1 of pertactin.
SQ	Sequence 15 AA:
Query Match	100.0%; Score 27; DB 23; Length 15;
Best Local Similarity	100.0%; Pred. NO. 47;
Matches 5; Conservative	0; Mismatches 0; Indels 0; Gaps 0
QY	1 GGAVP 5
DB	6 GGAVP 10
RESULT 5	
AAE17162	
ID	AAE17162 standard; peptide; 15 AA.
XX	
AC	AAE17162;
XX	
DT	18-APR-2002 (first entry)
DE	Monoclonal Ab derived Pep3 peptide specific to region 1 of pertactin.
XX	
KW	Pertactin; prnl; antibacterial; immunostimulant; antimicrobial; vaccine; diphtheria; tetanus; polio; Haemophilus influenza b infection; therapy; immune response.
XX	
OS	Unidentified.
XX	
PN	MO200200695-A2.

```

PD   03-JAN-2002.
XX
XX PF   29-JUN-2001; 2001WO-NL00493.
XX PR   30-JUN-2000; 2000EP-0202309.
XX PA      (NEME-) NEDERLANDEN MIN WELZIJN.
XX PI Mool FR;
XX DR WPI: 2002-139897/18.
XX
XX New polypeptides derived from Bordetella pertussis pertactin, useful as
PT a vaccine against infections caused by Bordetella strains, and other
PR infectious diseases of mammals, e.g. diphtheria, tetanus, or polio
PS Example 1; Fig 2b; 52pp; English.
XX
XX The invention relates to polypeptides derived from Bordetella pertussis
CC pertactin (Prnt). The polypeptide is useful in the preparation of
CC vaccines against B. pertussis, B. parapertussis, B. bronchiseptica and
CC other infectious diseases of mammals including diphtheria, tetanus, polio
CC and infections caused by Haemophilus influenza b. The polypeptide is
CC especially useful for eliciting an immune response against Bordetella sp.
CC Antibodies against the polypeptide may be used for pharmaceutical and/or
CC diagnostic purposes, particularly for treating or preventing infections
CC caused by Bordetella pertussis or Bordetella parapertussis. The present
CC sequence is a peptide derived from monoclonal antibody which is specific
CC to region I of pertactin.
XX
XX Sequence 15 AA:
SQ
Query Match          100.0%; Score 27; DB 23; Length 15;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
OY    1 GGAVP 5
       . | | | |
DB     1 GGAVP 5
RESULT 6
AAEI7163
ID AAEI7163 standard; peptide; 15 AA.
XX AC AAEI7163;
XX DT 18-APR-2002 (first entry)
XX DE Monoclonal Ab derived Pep4 peptide specific to region I of pertactin.
XX KW Pertactin; prnl; antibacterial; immunostimulant; antimicrobial; vaccine;
KW diphtheria; tetanus; polio; Haemophilus influenza b infection; therapy;
XX immune response.
OS Unidentified.
XX WO200200695-A2.
XX PN 03-JAN-2002.
XX PD 29-JUN-2001; 2001WO-NL00493.
XX PE 30-JUN-2000; 2000EP-0202309.
XX PR      (NEME-) NEDERLANDEN MIN WELZIJN.
XX PA Mool FR;
XX PI Mool FR;
XX DR WPI: 2002-139897/18.
XX
XX New polypeptides derived from Bordetella pertussis pertactin, useful as
PT a vaccine against infections caused by Bordetella strains, and other
PR infectious diseases of mammals, e.g. diphtheria, tetanus, or polio
PS Example 1; Fig 2b; 52pp; English.
XX
XX The invention relates to polypeptides derived from Bordetella pertussis
CC pertactin (Prnt). The polypeptide is useful in the preparation of
CC vaccines against B. pertussis, B. parapertussis, B. bronchiseptica and
CC other infectious diseases of mammals including diphtheria, tetanus, polio
CC and infections caused by Haemophilus influenza b. The polypeptide is
CC especially useful for eliciting an immune response against Bordetella sp.
CC Antibodies against the polypeptide may be used for pharmaceutical and/or
CC diagnostic purposes, particularly for treating or preventing infections
CC caused by Bordetella pertussis or Bordetella parapertussis. The present
CC sequence is a peptide derived from monoclonal antibody which is specific
CC to region I of pertactin.
XX
XX Sequence 15 AA:
SQ

```

PT Infectious diseases of mammals, e.g. diphtheria, tetanus, or polio
 XX
 PS Example 1; Fig 2b; 52pp; English.
 CC The invention relates to polypeptides derived from Bordetella pertussis
 CC pertactin (Pnrl). The polypeptide is useful in the preparation of
 CC vaccines against B. pertussis, B. parapertussis, B. bronchiseptica and
 CC other infectious diseases of mammals including diphtheria, tetanus, polio
 CC and infections caused by Haemophilus influenza b. The polypeptide is
 CC especially useful for eliciting an immune response against Bordetella sp.
 CC antibodies against the polypeptide may be used for pharmaceutical and/or
 CC diagnostic purposes, particularly for treating or preventing infections
 CC caused by Bordetella pertussis or Bordetella parapertussis. The present
 CC sequence is a peptide derived from monoclonal antibody which is specific
 CC to region 1 of pertactin.
 SQ Sequence 15 AA;
 Query Match 100.0%; Score 27; DB 23; Length 15;
 Best Local Similarity 100.0%; Pred. No. 47;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGAVP 5
 1 11111
 Db 1 GGAVP 5
 RESULT 7
 AAE17144
 ID AAE17144 standard; peptide; 20 AA.
 XX
 AC AAE17144;
 XX
 DT 18-APR-2002 (first entry)
 XX
 DE Bordetella pertussis pertactin (Pnrl) related peptide #3.
 XX
 KW Pertactin: pnr1; antibacterial; immunostimulant; antimicrobial; vaccine;
 KW diphtheria; tetanus; polio; Haemophilus influenza b infection; therapy;
 KW immune response.
 XX
 OS Undenitrified.
 XX
 PN WO200200695-A2.
 XX
 PD 03-JAN-2002.
 XX
 PF 29-JUN-2001; 2001WO-NL00493.
 XX
 PR 30-JUN-2000; 2000EP-0202309.
 XX
 PA (NEMO-) NEDERLANDEN MIN WEIZIJN.
 XX
 PI Wool FR;
 XX
 DR WPI: 2002-139897/18.
 XX
 PT New polypeptides derived from Bordetella pertussis pertactin, useful as
 PT a vaccine against infections caused by Bordetella strains, and other
 PT infectious diseases of mammals, e.g. diphtheria, tetanus, or polio
 PS Disclosure; Page 7; 52pp; English.
 XX
 CC The invention relates to polypeptides derived from Bordetella pertussis
 CC pertactin (Pnrl). The polypeptide is useful in the preparation of
 CC vaccines against B. pertussis, B. parapertussis, B. bronchiseptica and
 CC other infectious diseases of mammals including diphtheria, tetanus, polio
 CC and infections caused by Haemophilus influenza b. The polypeptide is
 CC especially useful for eliciting an immune response against Bordetella sp.
 CC antibodies against the polypeptide may be used for pharmaceutical and/or
 CC diagnostic purposes, particularly for treating or preventing infections
 CC caused by Bordetella pertussis or Bordetella parapertussis. The present
 CC sequence is B. pertussis pnr1 related peptide used in the invention.

XX
 SQ Sequence 20 AA;
 Query Match 100.0%; Score 27; DB 23; Length 20;
 Best Local Similarity 100.0%; Pred. No. 61;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGAVP 5
 1 11111
 Db 1 GGAVP 5
 RESULT 8
 AAU89387
 ID AAU89387 standard; peptide; 23 AA.
 XX
 AC AAU89387;
 XX
 DT 18-JUN-2002 (first entry)
 XX
 DE Insulin/insulin-like growth factor receptor-binding peptide #1343.
 XX
 KW Cytostatic; antidiabetic; neuroprotective; cerebroprotective;
 KW ophthalmological; insulin; receptor; gene therapy; diabetes;
 KW insulin-like growth factor-1; IGF-1; tumour; prostate; breast;
 KW diabetic retinopathy; neurological diseases; stroke;
 KW diabetic neuropathy.
 XX
 OS Synthetic.
 XX
 PN WO200172771-A2.
 XX
 PD 04-OCT-2001.
 XX
 PF 29-MAR-2000; 2000WO-US08528.
 XX
 PR 29-MAR-2000; 2000WO-US08528.
 XX
 PA (DGIB-) DGI BIOTECHNOLOGIES LLC.
 PA (NOVO) NOVO NORDISK AS.
 XX
 PI Beasley J, Blume AJ, Schaeffer L, Pillucia R, Brandt J;
 PI Brissette R, Spetzler J, Cheng W, Ostergaard S, Mandeckl WS;
 PI Hansen PH, Raveira M, Hsiao K;
 XX
 DR WPI: 2002-025774/03.
 XX
 PT Modulating insulin activity in mammalian cells, for treating e.g.
 PT diabetes and tumours, comprises using peptides that bind to insulin or
 PT insulin-like growth factor receptors
 PS Disclosure; Figure 2H-2; 390pp; English.
 XX
 CC The invention relates to a method of modulating insulin activity in
 CC mammalian cells by administering a peptide that binds the insulin
 CC receptor (IR). A composition containing a peptide, optionally expressed
 CC from gene therapy vectors, that binds to site 1 of IR and an insulin
 CC agonist are useful for treating diabetes. Also, peptides that are
 CC antagonists of the insulin-like growth factor-1 (IGF-1) receptor are
 CC useful for treating insulin-like growth factor (IGF) sensitive tumours
 CC (e.g. of prostate and breast) and diabetic retinopathy, while IGF-1
 CC receptor agonists are useful for treating neurological diseases,
 CC including stroke and diabetic neuropathy. The peptides are also useful in
 CC screening for compounds that bind to IR or IGF-1 receptor, potential
 CC therapeutics and research reagents. AAU8034-AAU90957 represent IR
 CC and/or IGF-1 receptor-binding peptides and related amino acid sequences
 CC of the invention.
 XX
 SQ Sequence 23 AA;
 Query Match 100.0%; Score 27; DB 23; Length 23;
 Best Local Similarity 100.0%; Pred. No. 70;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAVP 5
 DB 2 GGAVP 6

RESULT 9
 AAE17173
 ID AAE17173 standard; Protein: 31 AA.
 AC AAE17173;
 DE 18-APR-2002 (first entry)

Maltose binding protein-pertactin region 1 fusion protein, MBP-Prt4.
 Pertactin; prn1; antibacterial; immunostimulant; antimicrobial; vaccine;
 diphtheria; tetanus; polio; Haemophilus influenza b infection; therapy;
 immune response; maltose binding protein; MBP; fusion protein.

OS Chimeric - Bordetella pertussis.
 OS Chimeric - Unidentified.
 PN WO200200695-A2.
 PD 03-JAN-2002.
 PF 29-JUN-2001; 2001WO-NL00493.
 PR 30-JUN-2000; 2000EP-0202309.
 PA (NEWE-) NEDERLANDEN MIN WELZIJN.
 PI MOOI FR:
 WP1; 2002-139897/18.

New polypeptides derived from Bordetella pertussis pertactin, useful as
 a vaccine against infections caused by Bordetella strains, and other
 infectious diseases of mammals, e.g. diphtheria, tetanus, or polio -

Example 1; Fig 4b; 52pp; English.

The invention relates to polypeptides derived from Bordetella pertussis
 pertactin (Prn1). The polypeptide is useful in the preparation of
 vaccines against B. pertussis, B. parapertussis, B. bronchiseptica and
 other infectious diseases of mammals including diphtheria, tetanus, polio
 and infections caused by Haemophilus influenza b. The polypeptide is
 especially useful for eliciting an immune response against Bordetella sp.
 Antibodies against the polypeptide may be used for pharmaceutical and/or
 diagnostic purposes, particularly for treating or preventing infections
 caused by Bordetella pertussis or Bordetella parapertussis. The present
 sequence is maltose binding protein (MBP)-pertactin region 1 fusion
 protein used in the exemplification of the invention.

Sequence 31 AA;
 SQ

Query Match 100.0%; Score 27; DB 23; Length 31;
 Best Local Similarity 100.0%; Pred. No. 92;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAVP 5
 DB 10 GGAVP 14

RESULT 10
 AAE17141
 ID AAE17141 standard; peptide; 36 AA.
 AC AAE17141;
 DE 18-APR-2002 (first entry)

XX XX
 DE Bordetella pertussis pertactin (Prn1) peptide.
 XX Pertactin; prn1; antibacterial; immunostimulant; antimicrobial; vaccine;
 KW diphtheria; tetanus; polio; Haemophilus influenza b infection; therapy;
 KW immune response.
 XX Bordetella pertussis.
 OS WO200200695-A2.
 PN 03-JAN-2002.
 PD 29-JUN-2001; 2001WO-NL00493.
 PF 30-JUN-2000; 2000EP-0202309.
 PR (NEWE-) NEDERLANDEN MIN WELZIJN.
 PA MOOI FR:
 PI WP1; 2002-139897/18.
 DR

New polypeptides derived from Bordetella pertussis pertactin, useful as
 a vaccine against infections caused by Bordetella strains, and other
 infectious diseases of mammals, e.g. diphtheria, tetanus, or polio -

Claim 3; Page 4; 52pp; English.

The invention relates to polypeptides derived from Bordetella pertussis
 pertactin (Prn1). The polypeptide is useful in the preparation of
 vaccines against B. pertussis, B. parapertussis, B. bronchiseptica and
 other infectious diseases of mammals including diphtheria, tetanus, polio
 and infections caused by Haemophilus influenza b. The polypeptide is
 especially useful for eliciting an immune response against Bordetella sp.
 Antibodies against the polypeptide may be used for pharmaceutical and/or
 diagnostic purposes, particularly for treating or preventing infections
 caused by Bordetella pertussis or Bordetella parapertussis. The present
 sequence is B. pertussis prn1 peptide.

Sequence 36 AA;
 SQ

Query Match 100.0%; Score 27; DB 23; Length 36;
 Best Local Similarity 100.0%; Pred. No. 11e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAVP 5
 DB 10 GGAVP 14

RESULT 11
 AAE17170
 ID AAE17170 standard; Protein: 36 AA.
 AC AAE17170;
 DE 18-APR-2002 (first entry)

Maltose binding protein-pertactin region 1 fusion protein, MBP-Prn1.
 Pertactin; prn1; antibacterial; immunostimulant; antimicrobial; vaccine;
 KW diphtheria; tetanus; polio; Haemophilus influenza b infection; therapy;
 KW immune response; maltose binding protein; MBP; fusion protein.

OS Chimeric - Bordetella pertussis.
 OS Chimeric - Unidentified.
 PN WO200200695-A2.
 PD 03-JAN-2002.
 PF 29-JUN-2001; 2001WO-NL00493.

```

XX 30-JUN-2000; 2000EP-0202309.
XX (NEWE-) NEDERLANDEN MIN WELZIJN.
XX Mool FR:
XX WPI; 2002-139897/18.
XX
XX New polypeptides derived from Bordetella pertussis pertactin, useful as
XX a vaccine against infections caused by Bordetella strains, and other
XX infectious diseases of mammals, e.g. diphtheria, tetanus, or polio -
XX
XX Example 1; Fig 4b; 52pp; English.
XX
XX The invention relates to polypeptides derived from Bordetella pertussis
XX pertactin (Prnl). The polypeptide is useful in the preparation of
XX vaccines against B. pertussis, B. parapertussis, B. bronchiseptica and
XX other infectious diseases of mammals including diphtheria, tetanus, polio
XX and infectious diseases caused by Haemophilus influenza b. The polypeptide is
XX especially useful for eliciting an immune response against Bordetella sp.
XX Antibodies against the polypeptide may be used for pharmaceutical and/or
XX diagnostic purposes, particularly for treating or preventing infections
XX caused by Bordetella pertussis or Bordetella parapertussis. The present
XX sequence is maltose binding protein (MBP)-pertactin region 1 fusion
XX protein used in the exemplification of the invention.
XX
XX Sequence 36 AA:
XX
XX Query Match 100.0%; Score 27; DB 23; Length 36;
XX Best Local Similarity 100.0%; Pred. No. 1.1e+02;
XX Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 GGAVP 5
XX |||||
XX Db 10 GGAVP 14
XX
XX RESULT 12
XX AAE17172
XX ID AAE17172 standard; Protein; 36 AA.
XX AC AAE17172;
XX AA 18-APR-2002 (first entry)
XX DE Maltose binding protein-pertactin region 1 fusion protein, MBP-Prn3.
XX KM Pertactin; prnl; antibacterial; immunostimulant; antimicrobial; vaccine;
XX KW diphtheria; tetanus; polio; Haemophilus influenza b infection; therapy;
XX KM immune response; maltose binding protein; MBP; fusion protein.
XX OS Chimeric - Bordetella pertussis.
XX OS Chimeric - unidentified.
XX PN WO200200695-A2.
XX XX
XX PD 03-JAN-2002.
XX PE 29-JUN-2001; 2001WO-NL00493.
XX PR 30-JUN-2000; 2000EP-0202309.
XX XX
XX PA (NEWE-) NEDERLANDEN MIN WELZIJN.
XX Mool FR:
XX WPI; 2002-139897/18.
XX
XX New polypeptides derived from Bordetella pertussis pertactin, useful as
XX a vaccine against infections caused by Bordetella strains, and other
XX infectious diseases of mammals, e.g. diphtheria, tetanus, or polio -
XX

```

```

PS Example 1; Fig 4b; 52pp; English.
XX
XX The invention relates to polypeptides derived from Bordetella pertussis
XX pertactin (Prnl). The polypeptide is useful in the preparation of
XX vaccines against B. pertussis, B. parapertussis, B. bronchiseptica and
XX other infectious diseases of mammals including diphtheria, tetanus, polio
XX and infectious diseases caused by Haemophilus influenza b. The polypeptide is
XX especially useful for eliciting an immune response against Bordetella sp.
XX Antibodies against the polypeptide may be used for pharmaceutical and/or
XX diagnostic purposes, particularly for treating or preventing infections
XX caused by Bordetella pertussis or Bordetella parapertussis. The present
XX sequence is maltose binding protein (MBP)-pertactin region 1 fusion
XX protein used in the exemplification of the invention.
XX
XX Sequence 36 AA:
XX
XX Query Match 100.0%; Score 27; DB 23; Length 36;
XX Best Local Similarity 100.0%; Pred. No. 1.1e+02;
XX Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 GGAVP 5
XX |||||
XX Db 10 GGAVP 14
XX
XX RESULT 13
XX AAE17174
XX ID AAE17174 standard; Protein; 36 AA.
XX AC AAE17174;
XX AA 18-APR-2002 (first entry)
XX DE Maltose binding protein-pertactin region 1 fusion protein, MBP-Prn5.
XX KM Pertactin; prnl; antibacterial; immunostimulant; antimicrobial; vaccine;
XX KW diphtheria; tetanus; polio; Haemophilus influenza b infection; therapy;
XX KM immune response; maltose binding protein; MBP; fusion protein.
XX OS Chimeric - Bordetella pertussis.
XX OS Chimeric - unidentified.
XX PN WO200200695-A2.
XX XX
XX PD 03-JAN-2002.
XX PE 29-JUN-2001; 2001WO-NL00493.
XX PR 30-JUN-2000; 2000EP-0202309.
XX XX
XX PA (NEWE-) NEDERLANDEN MIN WELZIJN.
XX Mool FR:
XX WPI; 2002-139897/18.
XX
XX New polypeptides derived from Bordetella pertussis pertactin, useful as
XX a vaccine against infections caused by Bordetella strains, and other
XX infectious diseases of mammals, e.g. diphtheria, tetanus, or polio -
XX
XX Example 1; Fig 4b; 52pp; English.
XX
XX The invention relates to polypeptides derived from Bordetella pertussis
XX pertactin (Prnl). The polypeptide is useful in the preparation of
XX vaccines against B. pertussis, B. parapertussis, B. bronchiseptica and
XX other infectious diseases of mammals including diphtheria, tetanus, polio
XX and infectious diseases caused by Haemophilus influenza b. The polypeptide is
XX especially useful for eliciting an immune response against Bordetella sp.
XX Antibodies against the polypeptide may be used for pharmaceutical and/or
XX diagnostic purposes, particularly for treating or preventing infections
XX caused by Bordetella pertussis or Bordetella parapertussis. The present
XX sequence is maltose binding protein (MBP)-pertactin region 1 fusion
XX protein used in the exemplification of the invention.
XX

```


XX Sequence 36 AA:

Query Match 100.0%; Score 27; DB 23; Length 36;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGAVP 5
| | | | |
DB 15 GGAVP 19

RESULT 14

ID AAR56849 standard; peptide: 38 AA.

AC AAR56849;

DT 13-FEB-1995 (first entry)

DE Lipid transfer protein P1.

KW Lipid transfer protein; LTP; differentiation; cell culture;

KM somatic plant embryo; embryogenesis; organogenesis; vine.

OS Vitis vinifera cv. Chasselas x Vitis berlandieri, line 41 B.

XX W09413787-A.

PD 23-JUN-1994.

PF 14-DEC-1993; 93WO-FR01239.

PR 14-DEC-1992; 92FR-0015044.

XX (LVMH-) LVMH RECH.

PI Boulay MPH, Coutos-Thévenot PML, Deloire AJ;

DR Jouenne TG, Maes OCA, Guern JRD;

XX WPI; 1994-217867/26.

PT Inducing differentiation of cultured cells with lipid transfer
protein - esp to produce somatic plant embryos, also new

PS proteins, DNA and oligo-nucleotide probes.

XX Claim 15; Fig 1; 36pp; French.

CC Differentiation of cells in culture (partic. for tissues or for
embryogenesis and/or organogenesis) is promoted by adding to the
culture medium at least one lipid transfer protein (LTP) or its
analogue. LTP are used to produce somatic plant embryos from
somatic cell cultures in vitro, esp. of vine cells (including
transformed cells). They can also be used in other biological
systems where lipid transfer/differentiation is required. LTP can
induce differentiation even under otherwise non-permissive
conditions (presence of auxin). LTP is esp. one of the four
proteins given in AAR56849-52, and is esp. derived from the same
species as the cells being cultured. LTP from carrot, spinach,
millet, maize, barley or rape can also be used (see AAR56853-56).

XX Sequence 38 AA;

Query Match 100.0%; Score 27; DB 15; Length 38;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGAVP 5
| | | | |
DB 21 GGAVP 25

RESULT 15

AAR56850
ID AAR56850 standard; peptide: 38 AA.

AC AAR56850;

DT 13-FEB-1995 (first entry)

DE Lipid transfer protein P2.

KW Lipid transfer protein; LTP; differentiation; cell culture;

KM somatic plant embryo; embryogenesis; organogenesis; vine.

OS Vitis vinifera cv. Chasselas x Vitis berlandieri, line 41 B.

XX W09413787-A.

PD 23-JUN-1994.

PF 14-DEC-1993; 93WO-FR01239.

PR 14-DEC-1992; 92FR-0015044.

XX (LVMH-) LVMH RECH.

PI Boulay MPH, Coutos-Thévenot PML, Deloire AJ;

DR Jouenne TG, Maes OCA, Guern JRD;

XX WPI; 1994-217867/26.

PT Inducing differentiation of cultured cells with lipid transfer
protein - esp to produce somatic plant embryos, also new

PS proteins, DNA and oligo-nucleotide probes.

XX Claim 15; Fig 1; 36pp; French.

CC Differentiation of cells in culture (partic. for tissues or for
embryogenesis and/or organogenesis) is promoted by adding to the
culture medium at least one lipid transfer protein (LTP) or its
analogue. LTP are used to produce somatic plant embryos from
somatic cell cultures in vitro, esp. of vine cells (including
transformed cells). They can also be used in other biological
systems where lipid transfer/differentiation is required. LTP can
induce differentiation even under otherwise non-permissive
conditions (presence of auxin). LTP is esp. one of the four
proteins given in AAR56849-52, and is esp. derived from the same
species as the cells being cultured. LTP from carrot, spinach,
millet, maize, barley or rape can also be used (see AAR56853-56).

XX Sequence 38 AA;

Query Match 100.0%; Score 27; DB 15; Length 38;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGAVP 5
| | | | |
DB 21 GGAVP 25

Search completed: March 17, 2003, 12:08:56
Job time : 31.3333 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 17, 2003, 12:10:29 ; Search time 11 Seconds
(Without alignments)
20.951 Million cell updates/sec

Title: SEQ25-AV

Perfect score: 27

Sequence: 1 ggapv 5

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 199416 seqs, 46092074 residues

Total number of hits satisfying chosen parameters: 199416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEWL_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PTCT_NEWL_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEWL_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US07_NEWL_PUB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEWL_PUB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEWL_PUB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEWL_PUB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	27	100.0	46	9	US-09-855-754-8
2	27	100.0	51	9	US-09-855-754-7
3	27	100.0	51	9	US-09-855-754-13
4	27	100.0	56	9	US-09-855-754-9
5	27	100.0	56	9	US-09-855-754-10
6	27	100.0	56	9	US-09-855-754-12
7	27	100.0	61	9	US-09-855-754-11
8	27	100.0	96	9	US-09-950-933A-42
9	27	100.0	107	10	US-09-764-869-708
10	27	100.0	133	10	US-09-867-550-372
11	27	100.0	135	10	US-09-864-761-45449
12	27	100.0	167	9	US-09-975-719-361
13	27	100.0	179	10	US-09-867-550-1032
14	27	100.0	266	10	US-09-791-171-153
15	27	100.0	284	10	US-09-759-847-2
16	27	100.0	340	12	US-10-138-787-3
17	27	100.0	340	12	US-10-021-121-4
18	27	100.0	364	9	US-09-738-626-6869
19	27	100.0	370	9	US-09-738-626-6865

20	27	100.0	417	9	US-10-174-590-474	Sequence 474, App
21	27	100.0	417	9	US-10-176-758-474	Sequence 474, App
22	27	100.0	417	9	US-10-175-737-474	Sequence 474, App
23	27	100.0	417	9	US-10-173-706-474	Sequence 474, App
24	27	100.0	417	9	US-10-175-738-474	Sequence 474, App
25	27	100.0	417	9	US-10-175-752-474	Sequence 474, App
26	27	100.0	417	9	US-10-176-482-474	Sequence 474, App
27	27	100.0	417	9	US-10-176-757-474	Sequence 474, App
28	27	100.0	417	9	US-10-176-913-474	Sequence 474, App
29	27	100.0	417	9	US-10-180-552-474	Sequence 474, App
30	27	100.0	417	9	US-10-180-557-474	Sequence 474, App
31	27	100.0	417	9	US-10-173-700-474	Sequence 474, App
32	27	100.0	417	9	US-10-174-572-474	Sequence 474, App
33	27	100.0	417	9	US-10-174-579-474	Sequence 474, App
34	27	100.0	417	9	US-10-174-582-474	Sequence 474, App
35	27	100.0	417	9	US-10-174-588-474	Sequence 474, App
36	27	100.0	417	9	US-10-175-739-474	Sequence 474, App
37	27	100.0	417	9	US-10-175-740-474	Sequence 474, App
38	27	100.0	417	9	US-10-175-743-474	Sequence 474, App
39	27	100.0	417	9	US-10-176-488-474	Sequence 474, App
40	27	100.0	417	9	US-10-176-742-474	Sequence 474, App
41	27	100.0	417	9	US-10-176-747-474	Sequence 474, App
42	27	100.0	417	9	US-10-176-750-474	Sequence 474, App
43	27	100.0	417	9	US-10-176-985-474	Sequence 474, App
44	27	100.0	417	9	US-10-176-987-474	Sequence 474, App
45	27	100.0	417	9	US-10-176-991-474	Sequence 474, App

ALIGNMENTS

RESULT 1
US-09-855-754-8
Sequence 8, Application US/09855754
Publication No. US20020192237A1
GENERAL INFORMATION:
APPLICANT: BOURSAX-ENDE, CAROLINE
TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED
TITLE OF INVENTION: REGIONS OF PERFECT IN BORDETELLA PERTUSSIS,
TITLE OF INVENTION: BORDETELLA PARAPERTUSSIS, AND BORDETELLA
TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS
FILE REFERENCE: 03495-0206-00000
CURRENT APPLICATION NUMBER: US/09/855,754
CURRENT FILING DATE: 2001-09-10
PRIOR APPLICATION NUMBER: 60/206,969
PRIOR FILING DATE: 2000-05-25
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 8
LENGTH: 46
TYPE: PRT
ORGANISM: Bordetella bronchiseptica
US-09-855-754-8
Query Match 100.0%; Score 27; DB 9; Length 46;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 GGAVP 5
DB 13 GGAVP 17
RESULT 2
US-09-855-754-7
Sequence 7, Application US/09855754
Publication No. US20020192237A1
GENERAL INFORMATION:
APPLICANT: BOURSAX-ENDE, CAROLINE
TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED

```

; TITLE OF INVENTION: REGIONS OF PERTACTIN IN BORDETELLA PERTUSSIS,
; TITLE OF INVENTION: BORDETELLA PARAPERTUSSIS, AND BORDETELLA
; TITLE OF INVENTION: BRONCHISEPTICA, THEIR USE IN DIAGNOSTICS, AND IN
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS
; FILE REFERENCE: 03495-0206-00000
; CURRENT APPLICATION NUMBER: US/09/855,754
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/206,969
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 51
; TYPE: PRT
; ORGANISM: Bordetella bronchiseptica
; US-09-855-754-7

```

```

Query Match          100.0%; Score 27; DB 9; Length 51;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```
QY      1 GGAVP 5
        |||||
Db       13 GGAVP 17

```

```

RESULT 3
; US-09-855-754-13
; Sequence 13, Application US/09855754
; Publication No. US20020192237A1
; GENERAL INFORMATION:
; APPLICANT: BOURSAX-ENDE, CAROLINE
; APPLICANT: GUISSO-MACLOUF, NICOLE
; TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED
; TITLE OF INVENTION: REGIONS OF PERTACTIN IN BORDETELLA PERTUSSIS,
; TITLE OF INVENTION: BRONCHISEPTICA, THEIR USE IN DIAGNOSTICS, AND IN
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS
; FILE REFERENCE: 03495-0206-00000
; CURRENT APPLICATION NUMBER: US/09/855,754
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/206,969
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 51
; TYPE: PRT
; ORGANISM: Bordetella bronchiseptica
; US-09-855-754-13

```

```

Query Match          100.0%; Score 27; DB 9; Length 51;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```
QY      1 GGAVP 5
        |||||
Db       13 GGAVP 17

```

```

RESULT 4
; US-09-855-754-9
; Sequence 9, Application US/09855754
; Publication No. US20020192237A1
; GENERAL INFORMATION:
; APPLICANT: BOURSAX-ENDE, CAROLINE
; APPLICANT: GUISSO-MACLOUF, NICOLE
; TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED
; TITLE OF INVENTION: REGIONS OF PERTACTIN IN BORDETELLA PERTUSSIS,
; TITLE OF INVENTION: BRONCHISEPTICA, THEIR USE IN DIAGNOSTICS, AND IN
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS
; FILE REFERENCE: 03495-0206-00000

```

```

; CURRENT APPLICATION NUMBER: US/09/855,754
; CURRENT FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/206,969
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Bordetella bronchiseptica
; US-09-855-754-9

```

```

Query Match          100.0%; Score 27; DB 9; Length 56;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```
QY      1 GGAVP 5
        |||||
Db       18 GGAVP 22

```

```

RESULT 5
; US-09-855-754-10
; Sequence 10, Application US/09855754
; Publication No. US20020192237A1
; GENERAL INFORMATION:
; APPLICANT: BOURSAX-ENDE, CAROLINE
; APPLICANT: GUISSO-MACLOUF, NICOLE
; TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED
; TITLE OF INVENTION: REGIONS OF PERTACTIN IN BORDETELLA PERTUSSIS,
; TITLE OF INVENTION: BRONCHISEPTICA, THEIR USE IN DIAGNOSTICS, AND IN
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS
; FILE REFERENCE: 03495-0206-00000
; CURRENT APPLICATION NUMBER: US/09/855,754
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/206,969
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Bordetella bronchiseptica
; US-09-855-754-10

```

```

Query Match          100.0%; Score 27; DB 9; Length 56;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```
QY      1 GGAVP 5
        |||||
Db       13 GGAVP 17

```

```

RESULT 6
; US-09-855-754-12
; Sequence 12, Application US/09855754
; Publication No. US20020192237A1
; GENERAL INFORMATION:
; APPLICANT: BOURSAX-ENDE, CAROLINE
; APPLICANT: GUISSO-MACLOUF, NICOLE
; TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED
; TITLE OF INVENTION: REGIONS OF PERTACTIN IN BORDETELLA PERTUSSIS,
; TITLE OF INVENTION: BRONCHISEPTICA, THEIR USE IN DIAGNOSTICS, AND IN
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS
; FILE REFERENCE: 03495-0206-00000
; CURRENT APPLICATION NUMBER: US/09/855,754
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/206,969
; NUMBER OF SEQ ID NOS: 24

```

SOFTWARE: Patentln Ver. 2.1
 SEQ ID NO 12
 LENGTH: 56
 TYPE: PRT
 ORGANISM: Bordetella bronchiseptica
 US-09-855-754-12

Query Match 100.0%; Score 27; DB 9; Length 56;
 Best Local Similarity 100.0%; Pred. No. 30;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAVP 5
 |||||
 DB 13 GGAVP 17

RESULT 7
 US-09-855-754-11
 Sequence 11, Application US/09855754
 Publication No. US20020192237A1
 GENERAL INFORMATION:
 APPLICANT: BOUSSAUX-EUDE, CAROLINE
 APPLICANT: GUISSO-MACLOUP, NICOLE
 TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED
 TITLE OF INVENTION: REGIONS OF PERTACTIN IN BORDETELLA PERTUSSIS,
 TITLE OF INVENTION: BORDETELLA PARAPERTUSSIS, AND BORDETELLA
 TITLE OF INVENTION: BRONCHISEPTICA, THEIR USE IN DIAGNOSTICS, AND IN
 TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS
 FILE REFERENCE: 03495-0206-00000
 CURRENT APPLICATION NUMBER: US/09/855,754
 CURRENT FILING DATE: 2001-09-10
 PRIOR APPLICATION NUMBER: 60/206,969
 PRIOR FILING DATE: 2000-05-25
 NUMBER OF SEQ ID NOS: 24
 SOFTWARE: Patentln Ver. 2.1
 SEQ ID NO 11
 LENGTH: 61
 TYPE: PRT
 ORGANISM: Bordetella bronchiseptica
 US-09-855-754-11

Query Match 100.0%; Score 27; DB 9; Length 61;
 Best Local Similarity 100.0%; Pred. No. 33;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAVP 5
 |||||
 DB 13 GGAVP 17

RESULT 8
 US-09-950-933A-42
 Sequence 42, Application US/09950933A
 Patent No. US2002016141A1
 GENERAL INFORMATION:
 APPLICANT: Stimmoms, Carl R.
 APPLICANT: Navarro, Pedro
 TITLE OF INVENTION: Antimicrobial Peptides and Methods of
 TITLE OF INVENTION: Use
 FILE REFERENCE: 35718/238472
 CURRENT APPLICATION NUMBER: US/09/950,933A
 CURRENT FILING DATE: 2001-09-11
 PRIOR APPLICATION NUMBER: 60/232,569
 PRIOR FILING DATE: 2000-09-13
 NUMBER OF SEQ ID NOS: 99
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 42
 LENGTH: 96
 TYPE: PRT
 ORGANISM: Zea mays
 US-09-950-933A-42

Query Match 100.0%; Score 27; DB 9; Length 96;

Best Local Similarity 100.0%; Pred. No. 53;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGAVP 5
 |||||
 DB 29 GGAVP 33

RESULT 9
 US-09-764-869-708
 Sequence 708, Application US/09764869
 Patent No. US20020061521A1
 GENERAL INFORMATION:
 APPLICANT: Rosen et al.
 TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 FILE REFERENCE: PC007
 CURRENT APPLICATION NUMBER: US/09/764,869
 CURRENT FILING DATE: 2001-01-17
 Prior application data removed - refer to PALM or file wrapper
 NUMBER OF SEQ ID NOS: 2442
 SOFTWARE: Patentln Ver. 2.0
 SEQ ID NO 708
 LENGTH: 107
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: SITE
 LOCATION: (68)
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 NAME/KEY: SITE
 LOCATION: (75)
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 NAME/KEY: SITE
 LOCATION: (83)
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 US-09-764-869-708

Query Match 100.0%; Score 27; DB 10; Length 107;
 Best Local Similarity 100.0%; Pred. No. 59;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAVP 5
 |||||
 DB 27 GGAVP 31

RESULT 10
 US-09-867-550-372
 Sequence 372, Application US/09867550
 Patent No. US20020082206A1
 GENERAL INFORMATION:
 APPLICANT: Leach, Martin D.
 APPLICANT: Mehraban, Foad
 APPLICANT: Conley, Pamela
 APPLICANT: Law, Debbie
 APPLICANT: Topper, James
 TITLE OF INVENTION: No. US20020082206A1 Polynucleotides from Atherogenic Cells a
 FILE REFERENCE: 21402-013 (Cura-313)
 CURRENT APPLICATION NUMBER: US/09/867,550
 CURRENT FILING DATE: 2001-09-20
 PRIOR APPLICATION NUMBER: USSN 60/208,427
 PRIOR FILING DATE: 2000-05-30
 NUMBER OF SEQ ID NOS: 2125
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 372
 LENGTH: 133
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: VARIANT
 LOCATION: (118)
 OTHER INFORMATION: wherein Xaa may be any one of Asp or Glu

US-09-867-550-372

Query Match 100.0%; Score 27; DB 10; Length 133;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGAVP 5
DB 8 GGAVP 12

RESULT 11

US-09-864-761-45449
Sequence 45449, Application US/09864761
Patent No. US20020048763A1

GENERAL INFORMATION:

APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aemica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23

PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04

PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: GB 24263,6
PRIOR FILING DATE: 2000-10-04

PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27

PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21

PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30

PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29

NUMBER OF SEQ ID NOS: 49117

SOFTWARE: Annomax Sequence Listing Engine vers. 1.1

SEQ ID NO 45449

LENGTH: 135

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: MAP TO AL161422.3

OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1

OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.54

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.52

OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.63

OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.82
OTHER INFORMATION: SWISSPROT HIT: 009289, EVALU 5.90e+00
OTHER INFORMATION: EST_HUMAN HIT: W56629.1, EVALU 4.00e-39
US-09-864-761-45449

Query Match 100.0%; Score 27; DB 10; Length 135;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGAVP 5
DB 77 GGAVP 81

RESULT 12

US-09-975-719-361
Sequence 361, Application US/09975719
Publication No. US20030022349A1

GENERAL INFORMATION:

APPLICANT: Ausubel, Frederick M.
APPLICANT: Rahme, Laurence G.

TITLE OF INVENTION: VIRULENCE-ASSOCIATED NUCLEIC ACID
FILE REFERENCE: 00786/361003
CURRENT APPLICATION NUMBER: US/09/975,719
CURRENT FILING DATE: 2001-10-10

PRIOR APPLICATION NUMBER: US 09/199,637
PRIOR FILING DATE: 1998-11-25

PRIOR APPLICATION NUMBER: US 60/066,517
PRIOR FILING DATE: 1997-11-25

NUMBER OF SEQ ID NOS: 437
SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 361

LENGTH: 167

TYPE: PRT

ORGANISM: Pseudomonas aeruginosa

Query Match 100.0%; Score 27; DB 9; Length 167;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGAVP 5
DB 39 GGAVP 43

RESULT 13

US-09-867-550-1032
Sequence 1032, Application US/09867550
Patent No. US20020082206A1

GENERAL INFORMATION:

APPLICANT: Leach, Martin D.
APPLICANT: Mehraban, Foad,
APPLICANT: Conley, Pamela

APPLICANT: Law, Debbie
TITLE OF INVENTION: Topper, James

FILE REFERENCE: 21402-013 (Cura-313)
CURRENT APPLICATION NUMBER: US/09/867,550
CURRENT FILING DATE: 2001-09-20

PRIOR APPLICATION NUMBER: USSN 60/208,427
PRIOR FILING DATE: 2000-05-30

NUMBER OF SEQ ID NOS: 2125
SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 1032

LENGTH: 179

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: MAP TO AL161422.3

OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1

OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.54

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.52

OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.63

Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGAVP 5
Db 161 GGAVP 165

RESULT 14

US-09-791-171-153
Sequence 153, Application US/09791171
Patent No. US20020094336A1
GENERAL INFORMATION:
APPLICANT: ANDERSEN, Peter
APPLICANT: NIELSEN, Rikke
APPLICANT: OETTINGER, Thomas
APPLICANT: RASMUSSEN, Peter Birk
APPLICANT: ROSENKRANDS, Ida
APPLICANT: WEIDINGH, Karin
APPLICANT: FLORIO, Walter
TITLE OF INVENTION: NUCLEIC ACIDS FRAGMENTS AND POLYPEPTIDE FRAGMENTS
TITLE OF INVENTION: DERIVED FROM M. TUBERCULOSIS
FILE REFERENCE: 670001-2002.1
CURRENT APPLICATION NUMBER: US/09/791,171
CURRENT FILING DATE: 2001-02-20
PRIOR APPLICATION NUMBER: 09/050,739
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 0376/97
PRIOR FILING DATE: 1997-04-02
PRIOR APPLICATION NUMBER: 1277/97
PRIOR FILING DATE: 1997-11-10
PRIOR APPLICATION NUMBER: 60/044,624
PRIOR FILING DATE: 1997-04-18
PRIOR APPLICATION NUMBER: 60/070,488
PRIOR FILING DATE: 1998-01-05
NUMBER OF SEQ ID NOS: 173
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 153
LENGTH: 266
TYPE: PRT
ORGANISM: Mycobacterium tuberculosis
US-09-791-171-153

Query Match 100.0%; Score 27; DB 10; Length 266;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGAVP 5
Db 132 GGAVP 136

RESULT 15

US-09-759-847-2
Sequence 2, Application US/09759847
Patent No. US20020082410A1
GENERAL INFORMATION:
APPLICANT: Edlund, Thomas
TITLE OF INVENTION: Insulin Promoter Factor, and Uses
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Asclit (text)

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/759,847
FILING DATE: 12-Jan-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/031,898
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: ONI-004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 284 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-759-847-2

Query Match 100.0%; Score 27; DB 10; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGAVP 5
Db 246 GGAVP 250

Search completed: March 17, 2003, 12:21:35
Job time : 12 secs

10

11

12

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 17, 2003, 12:06:28 ; Search time 12.3333 Seconds
(without alignments) 38.973 Million cell updates/sec

Title: SEQ25-AV
Sequence: 1 ggapv 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	27	100.0	38	2 S39034	lipid transfer pro
2	27	100.0	92	2 S41723	hypothetical prote
3	27	100.0	100	2 S22567	hypothetical prote
4	27	100.0	100	2 S16054	hypothetical prote
5	27	100.0	107	2 E72710	hypothetical prote
6	27	100.0	116	2 S71564	lipid transfer pro
7	27	100.0	116	2 T03489	conserved hypotet
8	27	100.0	117	2 A12079	non-specific lipid
9	27	100.0	132	2 A95965	hypothetical expor
10	27	100.0	145	2 T22693	hypothetical prote
11	27	100.0	148	2 S46044	hypothetical prote
12	27	100.0	175	2 B69165	endonuclease III h
13	27	100.0	179	2 A33164	hypothetical prote
14	27	100.0	181	2 B70704	probable rocd - My
15	27	100.0	186	1 A09ENV	aqueous precursor
16	27	100.0	196	2 A26623	aqueous precursor
17	27	100.0	198	2 S39022	aqueous precursor
18	27	100.0	201	2 T35017	mitochondrion precurs
19	27	100.0	217	2 H69669	protocatechuate 3,
20	27	100.0	217	2 D69670	cholesterol ABC transp
21	27	100.0	218	2 E82780	glycine betaine/can
22	27	100.0	237	2 G72676	endonuclease III x
23	27	100.0	242	2 T37107	hypothetical prote
24	27	100.0	245	2 B75340	hypothetical prote
25	27	100.0	254	2 A96928	transcriptional regu
26	27	100.0	257	2 G82951	transcriptional re
27	27	100.0	259	2 H83920	probable transcript
28	27	100.0	266	2 H70858	transcription regu
29	27	100.0	272	1 LPEC28	probable electron
					lipoprotein-28 pre

30	27	100.0	272	2 C91203	lipoprotein-28 [im
31	27	100.0	272	2 D86049	lipoprotein-28 [im
32	27	100.0	274	2 AH2942	4-deoxy-L-threo-5-
33	27	100.0	277	2 B98340	5-keto-4-deoxyuron
34	27	100.0	283	2 B70683	probable cyst - My
35	27	100.0	283	2 G01926	insulin promoter f
36	27	100.0	283	2 B72232	ATP synthase F0, s
37	27	100.0	284	2 S39581	IPPI protein - mou
38	27	100.0	289	2 T02839	probable membrane
39	27	100.0	294	2 B70975	hypothetical prote
40	27	100.0	301	2 T35805	hypothetical prote
41	27	100.0	320	2 T36092	hypothetical prote
42	27	100.0	320	2 S32966	probable membrane
43	27	100.0	323	2 F95398	probable ABC trans
44	27	100.0	331	2 AB1323	ketol-acid reducto
45	27	100.0	331	2 AC1694	ketol-acid reducto

ALIGNMENTS

```

RESULT 1
S39034
lipid transfer protein p1 - grape
C:Species: Vitis sp. (grape)
C>Date: 25-Dec-1994 #sequence_revision 19-Apr-1996 #text_change 17-Mar-1999
C:Accession: S39034
R:Contos-Thevenot, P.; Jouenne, T.; Maes, O.; Guerbet, F.; Grosbois, M.; le Caer, J
Eur. J. Biochem. 217, 885-889, 1993
A:Title: Four 9-kDa proteins excreted by somatic embryos of grapevine are isoforms of
A:Reference number: S39034; MUID:94039144; PMID:8223644
A:Accession: S39034
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-38 <PGR>
C:Superfamily: phospholipid transfer protein

Query Match          100.0%  Score 27;  DB 2;  Length 38;
Best Local Similarity 100.0%  Pred. No. 25;
Matches 5;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

QY 1 GGAVP 5
    |||||
DB 21 GGAVP 25

RESULT 2
S41723
hypothetical protein - Saccharopolyspora erythraea
C:Species: Saccharopolyspora erythraea
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 22-Oct-1999
C:Accession: S41723
R:Brown, D.P.; Idler, K.B.; Backer, D.M.; Donadio, S.; Katz, L.
Mol. Gen. Genet. 242, 185-193, 1994
A:Title: Characterization of the genes and attachment sites for site-specific integras
A:Reference number: S41722; MUID:94211208; PMID:8159169
A:Accession: S41723
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-92 <BRO>
A:Cross-references: EMBL:L11597; NID:9404798; PIDN:AAA26478.1; PID:9404800

Query Match          100.0%  Score 27;  DB 2;  Length 92;
Best Local Similarity 100.0%  Pred. No. 59;
Matches 5;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

QY 1 GGAVP 5
    |||||
DB 17 GGAVP 21

RESULT 3
S22567

```

hypothetical protein 2 - phage phi-C31
 C:Species: phage phi-C31
 C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 10-Nov-1995
 C:Accession: S22567
 R:Rausch, H.; Lehmann, M.
 Nucleic Acids Res. 19, 5187-5189, 1991
 A:Title: Structural analysis of the actinophage phiC31 attachment site.
 A:Reference number: S22566; MUID:92020203; PMID:1656389
 A:Accession: S22567
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-100 <RAD>
 A:Cross-references: EMBL:X57036
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1990

Query Match 100.0%; Score 27; DB 2; Length 100;
 Best Local Similarity 100.0%; Pred. No. 64;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAVP 5
 |||||
 DB 86 GGAVP 90

RESULT 4
 S16054
 hypothetical protein 2 - phage phi-C31
 C:Species: phage phi-C31
 C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 08-Oct-1999
 C:Accession: S16054
 R:Rausch, H.
 submitted to the EMBL Data Library, December 1990
 A:Reference number: S16053
 A:Accession: S16054
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-100 <RAD>
 A:Cross-references: EMBL:X57036; NID:914810; PIDN:CAA40352.1; PID:9579069
 C:Genetics:
 A:Start codon: GTG

Query Match 100.0%; Score 27; DB 2; Length 100;
 Best Local Similarity 100.0%; Pred. No. 64;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAVP 5
 |||||
 DB 86 GGAVP 90

RESULT 5
 E72710
 hypothetical protein APE1100 - Aeropyrum pernix (strain K1)
 C:Species: Aeropyrum pernix
 C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
 C:Accession: E72710
 R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Hatakeyama, Y.; Jin-no, K.; Takahawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
 DNA Res. 6, 83-101, 1999
 A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyz
 A:Reference number: A72450; MUID:99310339; PMID:10382966
 A:Accession: E72710
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-107 <KAM>
 A:Cross-references: DDBJ:AP000060; NID:95104188; PIDN:BAA80085.1; PID:dl043871; PID:9510
 A:Experimental source: strain K1
 C:Genetics:
 A:Gene: APE1100

Query Match 100.0%; Score 27; DB 2; Length 107;
 Best Local Similarity 100.0%; Pred. No. 69;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAVP 5
 |||||
 DB 72 GGAVP 76

RESULT 6
 S71564
 lipid transfer protein SDI-9, drought-induced - common sunflower
 C:Species: Helianthus annuus (common sunflower)
 C:Date: 27-Nov-1997 #sequence_revision 12-Dec-1997 #text_change 07-May-1999
 C:Accession: S71564
 R:Duviols, O.; Cellier, F.; Ferrare, K.; Tausch, D.; Lamaze, T.; Dupuis, J.M.; Casse-
 Plant Mol. Biol. 31, 819-829, 1996
 A:Title: Identification and expression of water stress- and abscisic acid-regulated g
 A:Reference number: S71560; MUID:96400036; PMID:8806412
 A:Accession: S71564
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-116 <OUT>
 A:Cross-references: EMBL:X92648
 C:Superfamily: phospholipid transfer protein
 C:Keywords: lipid transport; stress-induced protein

Query Match 100.0%; Score 27; DB 2; Length 116;
 Best Local Similarity 100.0%; Pred. No. 75;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAVP 5
 |||||
 DB 45 GGAVP 49

RESULT 7
 T03489
 conserved hypothetical protein - Rhodobacter capsulatus
 C:Species: Rhodobacter capsulatus
 C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 08-Oct-1999
 C:Accession: T03489
 R:Vlcek, C.; Paces, V.; Maltsev, N.; Paces, J.; Haselkorn, R.; Fonstein, M.
 Proc. Natl. Acad. Sci. U.S.A. 94, 9384-9386, 1997
 A:Title: Sequence of a 189-kb segment of the chromosome of Rhodobacter capsulatus SB1
 A:Reference number: Z14955; MUID:97404404; PMID:9256491
 A:Accession: T03489
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-116 <VIC>
 A:Cross-references: EMBL:AF010496; NID:93128256; PIDN:AA016142.1; PID:93128290
 C:Genetics:
 A:Map position: 1

Query Match 100.0%; Score 27; DB 2; Length 116;
 Best Local Similarity 100.0%; Pred. No. 75;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAVP 5
 |||||
 DB 72 GGAVP 76

RESULT 8
 T12079
 non-specific lipid transfer protein LTP-24, drought and ABA induced - kidney bean
 C:Species: Phaseolus vulgaris (kidney bean)
 C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
 C:Accession: T12079
 R:Colmenero-Flores, J.M.; Campos, F.; Garciaarrubio, A.; Covarrubias, A.A.
 Plant Mol. Biol. 35, 393-405, 1997
 A:Title: Characterization of Phaseolus vulgaris CDNA clones responsive to water defic
 A:Reference number: Z17404; MUID:98009976; PMID:9349263
 A:Accession: T12079
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: mRNA

A:Residues: 1-117 <COL>
 A:Cross-references: EMBL:U72765; NID:92347087; PIDN:AC49660.1; PID:92347088
 A:Experimental source: water deficit- and ABA-treated plants
 C:Genetics:
 A:Gene: ltp-24
 C:Superfamily: phospholipid transfer protein
 C:Keywords: lipid transport; stress-induced protein

Query Match 100.0%; Score 27; DB 2; Length 117;
 Best Local Similarity 100.0%; Pred. No. 75;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 GGAVP 5
 Db 46 GGAVP 50

RESULT 9

hypothetical exported protein [imported] - Sinorhizobium meliloti (strain 1021) magaplas
 C:Species: Sinorhizobium meliloti
 C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
 C:Accession: A95965
 R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmaster, J.; Chain, P.; Vorholter, F.J.; Hernan
 Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
 A:Title: The complete sequence of the 1.683-kb pSymb megaplasmid from the N2-fixing endo
 A:Reference number: A95842; MUID:21396506; PMID:11481431
 A:Accession: A95965
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-132 <KUR>
 A:Cross-references: GB:AL591985; PIDN:CAC49385.1; PID:915140871; GSPDB:GN00167
 A:Experimental source: strain 1021, megaplasmid pSymb
 R:Gallibert, F.; Finan, T.M.; Long, S.R.; Punler, A.; Abola, P.; Ampe, F.; Barloy-Hubier,
 P.; Chain, P.; Cowle, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
 L.; Hyman, R.W.; Jones, T.
 Science 293, 668-672, 2001
 A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
 heubolt, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
 A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
 A:Reference number: A96039; MUID:21368234; PMID:11474104
 A:Contents: annotation
 C:Genetics:
 A:Gene: SMD21544
 A:Genome: plasmid

Query Match 100.0%; Score 27; DB 2; Length 132;
 Best Local Similarity 100.0%; Pred. No. 85;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 GGAVP 5
 Db 41 GGAVP 45

RESULT 10

T22693
 hypothetical protein F55B11.4 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T22693
 R:Ainscough, R.
 submitted to the EMBL Data Library, December 1996
 A:Reference number: Z19601
 A:Accession: T22693
 A:Status: preliminary; translated from GB/EMBL/DDBT
 A:Molecule type: DNA
 A:Residues: 1-145 <RTL>
 A:Cross-references: EMBL:Z83318; PIDN:CA805900.1; GSPDB:GN00022; CESP:F55B11.4
 A:Experimental source: clone F55B11
 C:Genetics:
 A:Gene: CESP:F55B11.4
 A:Map position: 4

A:introns: 15/1; 54/1; 114/1

Query Match 100.0%; Score 27; DB 2; Length 145;
 Best Local Similarity 100.0%; Pred. No. 93;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 GGAVP 5
 Db 107 GGAVP 111

RESULT 11

S46044
 hypothetical protein YBR173c - yeast (Saccharomyces cerevisiae)
 N:Alternate names: hypothetical protein YBR1234
 C:Species: Saccharomyces cerevisiae
 C:Date: 26-Aug-1994 #sequence_revision 09-Sep-1994 #text_change 19-Apr-2002
 C:Accession: S46044
 R:Entian, K.D.; Koetter, P.; Rose, M.; Becker, J.; Grey, M.; Li, Z.; Niegemann, E.; S
 S.; Grunstein, R.; Hedges, D.; Klesau, P.; Korol, S.; Krebs, B.; Proft, M.; Slegers,
 submitted to the Protein Sequence Database, August 1994
 A:Reference number: S46013
 A:Accession: S46044
 A:Molecule type: DNA
 A:Residues: 1-148 <ENT>
 A:Cross-references: EMBL:Z36042; NID:9536519; PID:9536520; GSPDB:GN00002; MIPS:YBR173
 A:Experimental source: strain S288C
 C:Genetics:
 A:Gene: SGD:UMP1; MIPS:YBR173c
 A:Cross-references: SGD:S0000377
 A:Map position: 2R

Query Match 100.0%; Score 27; DB 2; Length 148;
 Best Local Similarity 100.0%; Pred. No. 95;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 GGAVP 5
 Db 38 GGAVP 42

RESULT 12

B69165
 endonuclease III homolog - Methanobacterium thermoautotrophicum (strain Delta H)
 C:Species: Methanobacterium thermoautotrophicum
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
 C:Accession: B69165
 R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T
 ; Qiu, D.; Spadafora, R.; Vicalre, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jivanli,
 J.; Bacteriol. 179, 7135-7155, 1997
 A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: fu
 A:Reference number: A69000; MUID:98037514; PMID:9371463
 A:Accession: B69165
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-175 <MTH>
 A:Cross-references: GB:AE000833; GB:AE000666; NID:92621559; PIDN:AA85002.1; PID:9262
 A:Experimental source: strain Delta H
 C:Genetics:
 A:Gene: MTH496

Query Match 100.0%; Score 27; DB 2; Length 175;
 Best Local Similarity 100.0%; Pred. No. 114+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 GGAVP 5
 Db 63 GGAVP 67

RESULT 13

A33164

hypothetical protein walter - chicken
 M:Alternate names: homeotic protein Hox-3.1 homolog; hypothetical 13.1k protein
 C:Species: Gallus gallus (chicken)
 C:Date: 28-Mar-1991 #sequence_revision 28-Mar-1991 #text_change 21-Jul-2000
 C:Accession: A33164; JCI395
 R:Goldberg, G.S.
 Submitted to the Protein Sequence Database, March 1991
 A:Reference number: A33164
 A:Accession: A33164
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-179 <GOL>
 R:Experimental source: strain Leghorn
 R:Goldberg, G.S.; Kaczmarek, W.
 Gene 121, 397-398, 1992
 A:Title: Sequence of a novel chicken genomic DNA fragment that hybridizes to the murine
 A:Reference number: JCI395; MUID:9307061; PMID:1359990
 A:Accession: JCI395
 A:Molecule type: DNA
 A:Residues: 1-124 <GOL>
 A:Cross-references: GB:M84354; NID:q211948; PIDN:AAA70193.1; PID:q211949

Query Match 100.0%; Score 27; DB 2; Length 179;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAVP 5
 |||||
 DB 116 GGAVP 120

RESULT 14
 B70704
 Probable rocd - Mycobacterium tuberculosis (strain H37RV)
 C:Species: Mycobacterium tuberculosis
 C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
 C:Accession: B70704
 R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Fellwell, T.; Gencles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998
 A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A:Reference number: A70500; MUID:98295987; PMID:9634230
 A:Accession: B70704
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-181 <COL>
 A:Cross-references: GB:Z79702; GB:AL123456; NID:g3261642; PIDN:CAB02069.1; PID:e264239;
 A:Experimental source: strain H37RV
 C:Genetics:

Query Match 100.0%; Score 27; DB 2; Length 181;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAVP 5
 |||||
 DB 37 GGAVP 41

RESULT 15
 A03FNV
 aequorin precursor - hydromedusa (Aequorea victoria)
 C:Species: Aequorea victoria
 C:Date: 28-May-1986 #sequence_revision 28-May-1986 #text_change 26-May-2000
 C:Accession: A03020
 R:Inouye, S.; Noguchi, M.; Sakaki, Y.; Takagi, Y.; Miyata, T.; Iwanaga, S.; Miyata, T.; Proc. Natl. Acad. Sci. U.S.A. 82, 3154-3158, 1985
 A:Title: Cloning and sequence analysis of cDNA for the luminescent protein aequorin.
 A:Reference number: A03020; MUID:85216460; PMID:3858813
 A:Accession: A03020

A:Molecule type: mRNA
 A:Residues: 1-196 <INOS>
 A:Cross-references: GB:M11394; NID:g155658; PIDN:AAA27719.1; PID:g155659; GB:L29571;
 C:Experimental source: clone A0440
 C:Comment: The precise function of residues 1-7 is not known.
 C:Comment: The authors suggest that there are three calcium-binding sites, and that r
 ts a blue light.
 C:Comment: Cysteine residues appear to be critical to the ability of this photoprotein
 n, and mercaptoethanol is necessary to regenerate aequorin from apoaequorin.
 C:Superfamily: calmodulin; calmodulin repeat homology
 C:Keywords: calcium binding; EF hand; luminescence
 F:8-196/Product: aequorin #status predicted <MAT>
 F:18-50/Domain: calmodulin repeat homology <EF1>
 F:111-143/Domain: calmodulin repeat homology <EF2>
 F:147-179/Domain: calmodulin repeat homology <EF3>
 F:31,33,35,37,42/Binding site: calcium (Asp, Asn, Asn, Lys, Glu) #status predicted
 F:124,126,128,130,135/Binding site: calcium (Asp, Asp, Asn, Ala, Glu) #status predict
 F:160,162,164,166,171/Binding site: calcium (Asp, Asp, Ser, Glu, Glu) #status predict

Query Match 100.0%; Score 27; DB 1; Length 196;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAVP 5
 |||||
 DB 192 GGAVP 196

Search completed: March 17, 2003, 12:11:12
 Job time: 16.3333 secs

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 17, 2003, 12:05:19 ; Search time 7 Seconds
(without alignments)
29.626 Million cell updates/sec

Title: SEQ25-AV
Perfect score: 27
Sequence: 1 ggapv 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt-40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result. No.	Score	Query Match Length	ID	Description
1	27	100.0	NR1_VITX	P80275 vitis sp. (
2	27	100.0	NR2_VITX	P33556 vitis sp. (
3	27	100.0	NR1_PUOAR	P81651 prunus arne
4	27	100.0	NR1_PUODO	P82534 prunus dome
5	27	100.0	NR1_PUPE	P81402 prunus pers
6	27	100.0	NR1_PUPE	P81402 prunus pers
7	27	100.0	NR1_PUPE	P81402 prunus pers
8	27	100.0	NR1_PUPE	P81402 prunus pers
9	27	100.0	NR1_PUPE	P81402 prunus pers
10	27	100.0	NR1_PUPE	P81402 prunus pers
11	27	100.0	NR1_PUPE	P81402 prunus pers
12	27	100.0	NR1_PUPE	P81402 prunus pers
13	27	100.0	NR1_PUPE	P81402 prunus pers
14	27	100.0	NR1_PUPE	P81402 prunus pers
15	27	100.0	NR1_PUPE	P81402 prunus pers
16	27	100.0	NR1_PUPE	P81402 prunus pers
17	27	100.0	NR1_PUPE	P81402 prunus pers
18	27	100.0	NR1_PUPE	P81402 prunus pers
19	27	100.0	NR1_PUPE	P81402 prunus pers
20	27	100.0	NR1_PUPE	P81402 prunus pers
21	27	100.0	NR1_PUPE	P81402 prunus pers
22	27	100.0	NR1_PUPE	P81402 prunus pers
23	27	100.0	NR1_PUPE	P81402 prunus pers
24	27	100.0	NR1_PUPE	P81402 prunus pers
25	27	100.0	NR1_PUPE	P81402 prunus pers
26	27	100.0	NR1_PUPE	P81402 prunus pers
27	27	100.0	NR1_PUPE	P81402 prunus pers
28	27	100.0	NR1_PUPE	P81402 prunus pers
29	27	100.0	NR1_PUPE	P81402 prunus pers
30	27	100.0	NR1_PUPE	P81402 prunus pers
31	27	100.0	NR1_PUPE	P81402 prunus pers
32	27	100.0	NR1_PUPE	P81402 prunus pers
33	27	100.0	NR1_PUPE	P81402 prunus pers

34	27	100.0	391	1	SBCD_BACSU	P23479 bacillus su
35	27	100.0	417	1	PHOA_ASPNG	P34724 aspergillus
36	27	100.0	423	1	TRI9_HUMAN	O9ns68 homo sapien
37	27	100.0	427	1	DEOA_MYCTU	O53366 mycobacteri
38	27	100.0	446	1	SYE_AZOB	P45631 azobactin
39	27	100.0	483	1	SYE_SYNY3	O55778 synechocyst
40	27	100.0	509	1	DLDH_CANFA	P49819 canis fami
41	27	100.0	524	1	CCBI_RABIT	P19517 oryctolagus
42	27	100.0	548	1	ERF_HUMAN	P50548 homo sapien
43	27	100.0	551	1	ERF_MOUSE	P70459 mus musculu
44	27	100.0	557	1	ATKA_CLOAB	O32327 clostridium
45	27	100.0	559	1	FKB7_WHEAT	O43207 triticum ae

ALIGNMENTS

RESULT 1

ID	NR1_VITX	STANDARD:	PRT:	38 AA.
AC	P80275;			
DT	01-FEB-1994 (Rel. 28, Last sequence update)			
DT	01-FEB-1994 (Rel. 28, Last sequence update)			
DT	01-NOV-1997 (Rel. 35, Last annotation update)			
DE	Nonspecific lipid-transfer protein P1 (LTP P1) (Fragment).			
OS	vitis sp. (Grape).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Vitaceae;			
OC	Vitis.			
OX	NCBI_TaxID=3604;			
RN	[1]			
RP	SEQUENCE.			
RC	STRAIN-V. vinifera x Berlandieri cv. Rootstock 41B;			
RX	MDL:LINE-94039144; PubMed-8223644;			
RA	Coutos-Thievenot P., Jouenne T., Maes O., Guerbet F., Grosbois M.,			
RA	le Caer J.P., Boulay M., Deloire A., Kader J.-C., Guern J.;			
RT	"Four 9-kDa proteins excreted by somatic embryos of grapevine are			
RT	isoforms of lipid-transfer proteins."			
RI	Eur. J. Biochem. 217:885-889(1993).			
CC	- FUNCTION: PLANT NONSPECIFIC LIPID-TRANSFER PROTEINS TRANSFER			
CC	PHOSPHOLIPIDS AS WELL AS GALACTOLIPIDS ACROSS MEMBRANES. MAY PLAY			
CC	A ROLE IN WAX OR CUTIN DEPOSITION IN THE CELL WALLS OF EXPANDING			
CC	EPIDERMAL CELLS AND CERTAIN SECRETORY TISSUES.			
CC	- SIMILARITY: BELONGS TO THE PLANT LTP FAMILY.			
DR	HSP: P19656; IMZM.			
DR	InterPro: IPR000528; Plant-LTP.			
DR	InterPro: IPR001768; TRY/AMY1_inhbr.			
DR	Plant: PF00234; TRY_alpha_1.			
DR	PROSITE: PS00597; PLANT_LTP; PARTIAL.			
KW	Lipid-binding; Transport; Multigene family.			
FT	NON_TER			
SO	SEQUENCE 38 AA; 3626 MW; B5078F338325656 CRC64;			
Query Match 100.0%; Score 27; DB 1; Length 38;				
Best local Similarity 100.0%; Pred. No. 12;				
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
OY	1 GGAAP 5			
DB	21 GGAAP 25			
RESULT 2				
ID	NR2_VITX	STANDARD:	PRT:	38 AA.
AC	P33556;			
DT	01-FEB-1994 (Rel. 28, Created)			
DT	01-FEB-1994 (Rel. 28, Last sequence update)			
DT	01-NOV-1997 (Rel. 35, Last annotation update)			
DE	Nonspecific lipid-transfer protein P2 (LTP P2) (Fragment).			
OS	vitis sp. (Grape).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Vitaceae;			

OC Vitis.
 OX NCBI_TaxID=3604;
 RN [1]
 RP SEQUENCE.
 RC STRAIN-V. vinifera X Berlanthen cv. Rootstock 41B;
 RX MEDLINE=94039144; PubMed=8223644;
 RA Coutos-Thévenot P., Jousene T., Maes O., Guerbette F., Grosbois M.,
 le Caer J.P., Boulay M., Deloire A., Kader J.-C., Guern J.;
 RT "Four 9-kDa proteins secreted by somatic embryos of grapevine are
 isoforms of lipid-transfer proteins.";
 RL Eur. J. Biochem. 217:885-889(1993).
 CC -1- FUNCTION: PLANT NONSPECIFIC LIPID-TRANSFER PROTEINS TRANSFER
 PHOSPHOLIPIDS AS WELL AS GALACTOLIPIDS ACROSS MEMBRANES. MAY PLAY
 A ROLE IN WAX OR CUTIN DEPOSITION IN THE CELL WALLS OF EXPANDING
 EPIDERMAL CELLS AND CERTAIN SECRETORY TISSUES.
 CC -1- SIMILARITY: BELONGS TO THE PLANT LTP FAMILY.
 DR HSSP; P19656; IMZM.
 DR InterPro; IPR000528; Plant_LTP.
 DR InterPro; IPR001768; try_alpha_aml1.
 DR Pfam; PF00234; try_alpha_aml1.
 DR PROSITE; PS00597; PLANT_LTP; PARTIAL.
 KW Lipid-binding; Transport; Multigene family.
 FT NON_TER
 SQ SEQUENCE 38 AA; 3626 MW; B5078F338325C656 CRC64;

Query Match 100.0%; Score 27; DB 1; Length 38;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGAVP 5
 |||||
 DB - 21 GGAVP 25

RESULT 3
 ID NLT1_PRUAR STANDARD; PRT; 91 AA.
 AC P81651;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Nonspecific lipid-transfer protein 1 (LTP 1) (Major allergen Pru ar 3).
 OS Prunus armeniaca (Apricot).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eurosids I; Rosales; Rosaceae; Amygdaloideae; Prunus.
 OX NCBI_TaxID=36596;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=21312425; PubMed=11419703;
 RA Conti A., Fortunato D., Ortolani C., Giuffrida M.G., Pravettoni V.,
 Napolitano L., Farioli L., Perono Garoflo L., Trambaloli C.,
 Pastorello E.A.;
 RT "Determination of the primary structure of two lipid transfer proteins
 from apricot (Prunus armeniaca)";
 RL J. Chromatogr. B 756:123-129(2001).
 RN [2]
 RP SEQUENCE OF 1-34.
 RA Conti A., Baroglio C., Monza M., Ortolani C., Scibola E.,
 Giuffrida M.G., Fortunato D., Farioli L., Pravettoni V.,
 Napolitano L., Napolitano L., Pastorello E.A.;
 RT "The major allergen of apricot (Prunus armeniaca) is a lipid transfer
 protein.";
 RL Submitted (JAN-1999) to the SWISS-PROT data bank.
 CC -1- FUNCTION: PLANT NONSPECIFIC LIPID-TRANSFER PROTEINS TRANSFER
 PHOSPHOLIPIDS AS WELL AS GALACTOLIPIDS ACROSS MEMBRANES. MAY PLAY
 A ROLE IN WAX OR CUTIN DEPOSITION IN THE CELL WALLS OF EXPANDING
 EPIDERMAL CELLS AND CERTAIN SECRETORY TISSUES.
 CC -1- SIMILARITY: BELONGS TO THE PLANT LTP FAMILY.
 DR HSSP; P19656; IMZM.
 DR InterPro; IPR003612; AAI.
 DR InterPro; IPR000528; Plant_LTP.

DR InterPro; IPR001768; try_alpha_aml1_inhtr.
 DR Pfam; PF00234; try_alpha_aml1.
 DR PRINTS; PR00382; LIPIDTRANSFER.
 DR SMART; SM00499; AAI; 1.
 DR PROSITE; PS00597; PLANT_LTP; 1.
 KW Lipid-binding; Transport; Allergen.
 FT DISULFID 3 50 BY SIMILARITY.
 FT DISULFID 13 27 BY SIMILARITY.
 FT DISULFID 28 73 BY SIMILARITY.
 FT DISULFID 48 87 BY SIMILARITY.
 SQ SEQUENCE 91 AA; 9178 MW; 98A91F6B30CD8435 CRC64;

Query Match 100.0%; Score 27; DB 1; Length 91;
 Best Local Similarity 100.0%; Pred. No. 27;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGAVP 5
 |||||
 DB 20 GGAVP 24

RESULT 4
 ID NLT1_PRUDO STANDARD; PRT; 91 AA.
 AC P82534;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Nonspecific lipid-transfer protein 1 (LTP 1) (Major allergen Pru d 3).
 OS Prunus domestica (Garden plum).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eurosids I; Rosales; Rosaceae; Amygdaloideae; Prunus.
 OX NCBI_TaxID=3758;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Peelings;
 RX MEDLINE=21312422; PubMed=11419731;
 RA Pastorello E.A., Farioli L., Pravettoni V., Giuffrida M.G.,
 Ortolani C., Fortunato D., Trambaloli C., Scibola E., Calamari A.M.,
 Robino A.M., Conti A.;
 RT "Characterization of the major allergen of plum as a lipid transfer
 protein.";
 RL J. Chromatogr. B 756:95-103(2001).
 CC -1- FUNCTION: PLANT NONSPECIFIC LIPID-TRANSFER PROTEINS TRANSFER
 PHOSPHOLIPIDS AS WELL AS GALACTOLIPIDS ACROSS MEMBRANES. MAY PLAY
 A ROLE IN WAX OR CUTIN DEPOSITION IN THE CELL WALLS OF EXPANDING
 EPIDERMAL CELLS AND CERTAIN SECRETORY TISSUES. ALSO HAS AN
 IGE-BINDING ACTIVITY.
 CC -1- SIMILARITY: BELONGS TO THE PLANT LTP FAMILY.
 DR HSSP; P19656; IMZM.
 DR InterPro; IPR003612; AAI.
 DR InterPro; IPR000528; Plant_LTP.
 DR InterPro; IPR001768; try_alpha_aml1_inhtr.
 DR Pfam; PF00234; try_alpha_aml1.
 DR PRINTS; PR00382; LIPIDTRANSFER.
 DR SMART; SM00499; AAI; 1.
 DR PROSITE; PS00597; PLANT_LTP; 1.
 KW Lipid-binding; Transport; Allergen.
 FT DISULFID 3 50 BY SIMILARITY.
 FT DISULFID 13 27 BY SIMILARITY.
 FT DISULFID 28 73 BY SIMILARITY.
 FT DISULFID 48 87 BY SIMILARITY.
 SQ SEQUENCE 91 AA; 9174 MW; F2165CC465CD0377 CRC64;

Query Match 100.0%; Score 27; DB 1; Length 91;
 Best Local Similarity 100.0%; Pred. No. 27;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGAVP 5
 |||||
 DB 20 GGAVP 24

RESULT 5
NUTL_PROPE STANDARD: PRT; 91 AA.
AC P81402;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Nonspecific lipid-transfer protein 1 (LTP 1) (Major allergen Pru p 3)
DE (Pru p 1).
OS Prunus persica (Peach).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC euroids I; Rosales; Rosaceae; Amygdaloideae; Prunus.
OX NCBI_TaxID=3760;
RN [1]
RP SEQUENCE:
RX MEDLINE=20080487; PubMed=10614824;
RA Pastorello E.A., Farioli L., Pravelettoni V., Ortolani C., Espano M.,
RA Giuffrida M.G., Fortunato D., Farioli L., Monza M., Napolitano L.,
RA Sacco M., Scibola E., Conti A.,
RT "Complete amino acid sequence determination of the major allergen of
RT peach (Prunus persica) Pru p 1.";
RL Biol. Chem. 380:1315-1320(1999).
RN [2]
RP SEQUENCE OF 1-32: 53-66 AND 73-80.
RX MEDLINE=99169150; PubMed=10069889;
RA Pastorello E.A., Farioli L., Pravelettoni V., Ortolani C., Espano M.,
RA Monza M., Baroglio C., Scibola E., Ansaloni R., Incorvata C.,
RA Conti A.;
RT "The major allergen of peach (Prunus persica) is a lipid transfer
RT protein.";
RL J. Allergy Clin. Immunol. 103:520-526(1999).
CC -1- FUNCTION: PLANT NONSPECIFIC LIPID-TRANSFER PROTEIN. TRANSFER
CC PHOSPHOLIPIDS AS WELL AS GALACTOLIPIDS ACROSS MEMBRANES. MAY PLAY
CC A ROLE IN WAX OR CUTIN DEPOSITION IN THE CELL WALLS OF EXPANDING
CC EPIDERMAL CELLS AND CERTAIN SECRETORY TISSUES. ALSO HAS AN
CC IGE-BINDING ACTIVITY.
CC -1- SIMILARITY: BELONGS TO THE PLANT LTP FAMILY.
DR HSSP: P19656; IMZM.
DR InterPro: IPR003612; AAI.
DR InterPro: IPR000528; Plant_LTP.
DR InterPro: IPR001768; Try/amy1_inhbr.
DR Pfam: PF00234; try_alpha_amy1.1.
DR PRINTS: PR00382; LIPIDTRANSFER.
DR SMART: SM00499; AAI; 1.
DR PROSITE: PS00597; PLANT_LTP; 1.
KM Lipid-binding; Transport; Allergen.
FT DISULFID 3 50 BY SIMILARITY.
FT DISULFID 13 27 BY SIMILARITY.
FT DISULFID 28 73 BY SIMILARITY.
FT DISULFID 48 87 BY SIMILARITY.
SQ SEQUENCE 91 AA; 9178 MW; BB84569AA9E4B332 CRC64;
Query Match 100.0%; Score 27; DB 1; Length 91;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 GGAVP 5
DB 20 GGAVP 24
RESULT 6
NUTL_MALDO STANDARD: PRT; 115 AA.
AC 09M5X7; 09LEDO;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Nonspecific lipid-transfer protein precursor (LTP) (Allergen Mal d 3).
DE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC euroids I; Rosales; Rosaceae; Maloideae; Pyrus.
OX NCBI_TaxID=23211;
Malus domestica (Apple) (Malus sylvestris).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC euroids I; Rosales; Rosaceae; Maloideae; Malus.
OX NCBI_TaxID=3750;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-cv. Golden Delicious;
RA Schreier S., Mangorsch A., Hausteil D., Vieths S.;
RT "Cloning of a cDNA encoding a lipid transfer protein as a potential
RT allergen from Malus domestica.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 25-115 FROM N.A.
RA Diaz-Perales A., Garcia-Casado G., Sanchez-Monge R., Barber D.,
RA Salcedo G.;
RT "Cloning of a cDNA encoding a major allergen and lipid transfer
RT protein, from Malus domestica.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: PLANT NONSPECIFIC LIPID-TRANSFER PROTEIN. TRANSFER
CC PHOSPHOLIPIDS AS WELL AS GALACTOLIPIDS ACROSS MEMBRANES. MAY PLAY
CC A ROLE IN WAX OR CUTIN DEPOSITION IN THE CELL WALLS OF EXPANDING
CC EPIDERMAL CELLS AND CERTAIN SECRETORY TISSUES (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE PLANT LTP FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AF221502; AF26450.1; -;
DR EMBL: AJ277164; CAB96874.1; -;
DR HSSP: P19656; IMZM.
DR InterPro: IPR003612; AAI.
DR InterPro: IPR000528; Plant_LTP.
DR InterPro: IPR001768; Try/amy1_inhbr.
DR Pfam: PF00234; try_alpha_amy1.1.
DR PRINTS: PR00382; LIPIDTRANSFER.
DR SMART: SM00499; AAI; 1.
DR PROSITE: PS00597; PLANT_LTP; 1.
KM Lipid-binding; Transport; Signal; Allergen.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 115 NONSPECIFIC LIPID-TRANSFER PROTEIN.
FT DISULFID 27 74 BY SIMILARITY.
FT DISULFID 37 51 BY SIMILARITY.
FT DISULFID 52 97 BY SIMILARITY.
FT DISULFID 72 111 BY SIMILARITY.
SQ SEQUENCE 115 AA; 11410 MW; CF431D5A52AD8C95 CRC64;
Query Match 100.0%; Score 27; DB 1; Length 115;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 GGAVP 5
DB 44 GGAVP 48
RESULT 7
NUTL_PYRGO STANDARD: PRT; 115 AA.
AC 09M5X6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Nonspecific lipid-transfer protein precursor (LTP) (Allergen Pyr c 3).
DE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC euroids I; Rosales; Rosaceae; Maloideae; Pyrus.
OX NCBI_TaxID=23211;

```

RN [1]
RP SEQUENCE FROM N.A.
RA Schaefer S., Wagnorsch A., Hausteijn D., Vieths S.;
RT "Cloning of a cDNA encoding a lipid transfer protein as a potential
RT allergen from cyna communis."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: PLANT NONSPECIFIC LIPID-TRANSFER PROTEINS. TRANSFER
CC PHOSPHOLIPIDS AS WELL AS GALACTOLIPIDS ACROSS MEMBRANES. MAY PLAY
CC A ROLE IN WAX OR CUTIN DEPOSITION IN THE CELL WALLS OF EXPANDING
CC EPIDERMAL CELLS AND CERTAIN SECRETORY TISSUES (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE PLANT LTP FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; AF221503; AAF26451.1; -.
DR HSSP; P19656; IMZM.
DR InterPro; IPR003612; AAI.
DR InterPro; IPR000528; Plant_LTP.
DR InterPro; IPR001768; try/amyL_inhtr.
DR Pfam; PF00234; try_alpha_amyli; 1.
DR PRINTS; PR00382; LIPIDTRANSFER.
DR SMART; SM00499; AAI; 1.
DR PROSITE; PS00597; PLANT_LTP; 1.
DR LIPID-Binding; Transport; Signal; Allergen.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 115 NONSPECIFIC LIPID-TRANSFER PROTEIN.
FT DISULFID 27 74 BY SIMILARITY.
FT DISULFID 37 51 BY SIMILARITY.
FT DISULFID 52 97 BY SIMILARITY.
FT DISULFID 72 111 BY SIMILARITY.
SQ SEQUENCE 115 AA; 11463 MW; 90543BD036AA8D65 CMC64;

Query Match 100.0%; Score 27; DB 1; Length 115;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAVP 5
| | | | |
DB 44 GGAVP 48

RESULT 8
NLTLP_HELAN STANDARD; PRT; 116 AA.
AC Q39950;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DE Nonspecific lipid-transfer protein precursor (LTP) (NSLTP) (SDI-9).
OS Helianthus annuus (Common sunflower).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids II; Asterales; Asteraceae; Asteroideae;
OC Helianthaceae; Helianthus.
OX NCBI_TaxID=4232;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leaf;
RX MEDLINE=96400036; PubMed=8806412;
RA Ouyard O., Cellier F., Ferrare K., Tousch D., Lamaze T., Dupuis J.M.,
RA Casse-Delbart F.;
RT "Identification and expression of water stress- and abscisic acid-
RT regulated genes in a drought-tolerant sunflower genotype."
RL Plant Mol. Biol. 31:819-839(1996).
CC -1- FUNCTION: PLANT NONSPECIFIC LIPID-TRANSFER PROTEINS. TRANSFER
CC PHOSPHOLIPIDS AS WELL AS GALACTOLIPIDS ACROSS MEMBRANES. MAY PLAY
CC A ROLE IN WAX OR CUTIN DEPOSITION IN THE CELL WALLS OF EXPANDING

```

```

CC EPIDERMAL CELLS AND CERTAIN SECRETORY TISSUES.
CC -1- SIMILARITY: BELONGS TO THE PLANT LTP FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; X92648; CAA63340.1; -.
DR HSSP; P19656; IMZM.
DR InterPro; IPR003612; AAI.
DR InterPro; IPR000528; Plant_LTP.
DR InterPro; IPR001768; try/amyL_inhtr.
DR Pfam; PF00234; try_alpha_amyli; 1.
DR PRINTS; PR00382; LIPIDTRANSFER.
DR SMART; SM00499; AAI; 1.
DR LIPID-Binding; Transport; Signal.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 116 NONSPECIFIC LIPID-TRANSFER PROTEIN.
FT DISULFID 28 75 BY SIMILARITY.
FT DISULFID 38 52 BY SIMILARITY.
FT DISULFID 53 98 BY SIMILARITY.
FT DISULFID 73 112 BY SIMILARITY.
SQ SEQUENCE 116 AA; 11551 MW; 5AA38C96002F4B8 CMC64;

Query Match 100.0%; Score 27; DB 1; Length 116;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAVP 5
| | | | |
DB 45 GGAVP 49

RESULT 9
NLTLP_PRUDU STANDARD; PRT; 117 AA.
AC Q43017;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE Nonspecific lipid-transfer protein 1 precursor (LTP 1).
OS Prunus dulcis (Almond) (Prunus amygdalus).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Rosales; Rosaceae; Amygdaloideae; Prunus.
OX NCBI_TaxID=3755;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Texas; TISSUE=flower;
RA Esteban M.S.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: PLANT NONSPECIFIC LIPID-TRANSFER PROTEINS. TRANSFER
CC PHOSPHOLIPIDS AS WELL AS GALACTOLIPIDS ACROSS MEMBRANES. MAY PLAY
CC A ROLE IN WAX OR CUTIN DEPOSITION IN THE CELL WALLS OF EXPANDING
CC EPIDERMAL CELLS AND CERTAIN SECRETORY TISSUES.
CC -1- SIMILARITY: BELONGS TO THE PLANT LTP FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; X96714; CAA65475.1; -.
DR HSSP; P19656; IMZM.
DR InterPro; IPR003612; AAI.
DR InterPro; IPR000528; Plant_LTP.

```



```

DR InterPro: IPR001768; TRY/amyL inhbt.
DR Pfam: PF00234; tryp_alpha_amyL; 1.
DR PRINTS: PR00382; LIPIDTRANSFER.
DR SMART: SM00499; AAI; 1.
DR PROSITE: PS00597; PLANT_LTP; 1.
DR Lipid-binding; Transport; Signal; Multigene family.
FT SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 117 NONSPECIFIC LIPID-TRANSFER PROTEIN 1.
FT DISULFID 29 76 POTENTIAL.
FT DISULFID 39 53 POTENTIAL.
FT DISULFID 54 99 POTENTIAL.
FT DISULFID 74 113 POTENTIAL.
SQ SEQUENCE 117 AA; 11900 MW; 87533E117A11809C CRC64;

Query Match 100.0%; Score 27; DB 1; Length 117;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAVP 5
DB 46 GGAVP 50

RESULT 10
NLRP_PROAV STANDARD; PRT; 117 AA.
ID NLRP_PROAV
AC 09M5X8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Nonspecific lipid-transfer protein precursor (LTP) (Allergen Pru av
3).
OS Prunus avium (Cherry).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Rosales; Rosaceae; Amygdaloideae; Prunus.
OX NCBI_TaxID=42229;
RN [1]
RP SEQUENCE FROM N.A.
RA Scheurer S., Mangorsch A., Hausteiner D., Viethe S.;
RT "Cloning of a cDNA encoding a lipid transfer protein as a potential
RT allergen from Prunus avium."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: PLANT NONSPECIFIC LIPID-TRANSFER PROTEINS TRANSFER
CC PHOSPHOLIPIDS AS WELL AS GALACTOLIPIDS ACROSS MEMBRANES. MAY PLAY
CC A ROLE IN WAX OR CUTIN DEPOSITION IN THE CELL WALLS OF EXPANDING
CC EPIDERMAL CELLS AND CERTAIN SECRETORY TISSUES (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE PLANT LTP FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AF221501; AAF26449.1; -
DR HSSP: P19656; IMZM.
DR InterPro: IPR003612; AAI.
DR InterPro: IPR000528; PLANT_LTP.
DR InterPro: IPR001768; TRY/amyL_inhbt.
DR Pfam: PF00234; tryp_alpha_amyL; 1.
DR PRINTS: PR00382; LIPIDTRANSFER.
DR SMART: SM00499; AAI; 1.
DR PROSITE: PS00597; PLANT_LTP; 1.
DR Lipid-binding; Transport; Signal; Allergen.
FT SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 117 NONSPECIFIC LIPID-TRANSFER PROTEIN.
FT DISULFID 29 76 BY SIMILARITY.
FT DISULFID 39 53 BY SIMILARITY.
FT DISULFID 54 99 BY SIMILARITY.
FT DISULFID 74 113 BY SIMILARITY.

```

```

SQ SEQUENCE 117 AA; 11751 MW; 9E8DE1122205B48A CRC64;

Query Match 100.0%; Score 27; DB 1; Length 117;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAVP 5
DB 46 GGAVP 50

RESULT 11
UMPL_YEAST STANDARD; PRT; 148 AA.
ID UMLP_YEAST
AC P38293;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Proteasome maturation factor UMP1.
GN UMP1 OR YBR173C OR YBR1234.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-YH9500;
RX MEDLINE=96150854; PubMed=9491890;
RA Ramos P.C., Hoeckendorf J., Johnson E.S., Varshavsky A., Dohmen R.J.;
RT "Ump1 is required for proper maturation of the 20S proteasome and
RT becomes its substrate upon completion of the assembly."
RL Cell 92:489-499(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN-S288c;
RC Entian K.-D., Koetter P., Rose M., Becker J., Grey M., Li Z.,
RA Niegemann E., Schenk-Groeninger R., Servos J., Wehner E.,
RA Wolter R., Brendel M., Bauer J., Braun H., Derr K., Duesterhus S.,
RA Gruenbein R., Hedges D., Kleesau P., Korol S., Krebs B., Prott M.,
RA Siegers K., Baur A., Boles E., Miosga T.,
RA Schmitt-Gerstenschlaeger I., Zimmermann P.K.;
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: SHORT-LIVED CHAPERONE PRESENT IN THE PRECURSOR FORM OF
CC THE 20S PROTEASOME AND ABSENT IN THE MATURE COMPLEX. REQUIRED FOR
CC THE CORRECT ASSEMBLY AND ENZYMATIC ACTIVATION OF THE PROTEASOME.
CC ALSO PREVENTS PREMATURE PROCESSING OF THE PRE2 PROPEPTIDE.
CC -1- MUSCULANEUS: SEEMS TO BE DEGRADED BY THE PROTEASOME UPON ITS
CC FORMATION.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AJ002557; CA005556.1; -
DR EMBL: Z36042; CA085134.1; -
DR PIR: S46044; S46044.
DR SGD: S0000377; UMP1.
RW Proteasome; Chaperone.
SQ SEQUENCE 148 AA; 16760 MW; 7C774DDA0F3FCD7C CRC64;

Query Match 100.0%; Score 27; DB 1; Length 148;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAVP 5
DB 38 GGAVP 42

```

```

RESULT 12
AEQ1_AEQVI STANDARD; PRT; 196 AA.
ID AEQ1_AEQVI
AC P07164;
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Aequorin 1 precursor.
OS Aequorea victoria (Jellyfish).
OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydrozoa; Leptomedusae;
OC Aequoreidae; Aequorea.
OX NCBI_TaxId=6100;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87185437; PubMed=2882777;
RA Prasher D.C., McCann R.O., Longiaru M., Cormier M.J.;
RT "Sequence comparisons of complementary DNAs encoding aequorin
RT isotypes."
RL Biochemistry 26:1326-1332(1987).
RN [2]
RP SEQUENCE OF 8-196.
RX MEDLINE=86077721; PubMed=2866797;
RA Charbonneau H., Walsh K.A., McCann R.O., Prendergast F.G.,
RA Cormier M.J., Vanaman T.C.;
RT "Amino acid sequence of the calcium-dependent photoprotein aequorin."
RL Biochemistry 24:6762-6771(1985).
RN [3]
RP MUTAGENESIS OF PRO-196.
RX MEDLINE=92111761; PubMed=1765170;
RA Nomura M., Inouye S., Ohmura Y., Tsuji F.I.;
RT "A C-terminal proline is required for bioluminescence of the Ca(2+)-
RT binding photoprotein, aequorin."
RL FEBS Lett. 295:63-66(1991).
RN [4]
RP DISULFIDE BOND.
RX MEDLINE=94009705; PubMed=8405461;
RA Ohmura Y., Kurono S., Ohnishi M., Fagan T.F., Tsuji F.I.;
RT "Mass spectrometric evidence for a disulfide bond in aequorin
RT regeneration."
RL FEBS Lett. 332:226-228(1993).
RN [5]
RP FUNCTION: CA(2+)-DEPENDENT BIOLUMINESCENCE PHOTOPROTEIN. DISPLAYS
RN AN EMISSION PEAK AT 470 NM (BLUE LIGHT). TRACE AMOUNTS OF CALCIUM
RN ION TRIGGER THE INTRAMOLECULAR OXIDATION OF THE CHROMOPHORE,
RN COLENTERRAZINE INTO COLENTERRAMIDE AND CO(2) WITH THE
RN CONCOMITANT EMISSION OF LIGHT.
RN -1- PHE. THE REDUCTION OF THE DISULFIDE BOND IS NECESSARY TO
RN REGENERATE Aequorin FROM APOAequorin.
RN -1- SIMILARITY: BELONGS TO THE EF-HAND SUPERFAMILY. Aequorin FAMILY.
RN -1- SIMILARITY: CONTAINS 3 EF-HAND CALCIUM-BINDING DOMAINS.
RN -----
RN This SWISS-PROT entry is copyright. It is produced through a collaboration
RN between the Swiss Institute of Bioinformatics and the EMBL outstation -
RN the European Bioinformatics Institute. There are no restrictions on its
RN use by non-profit institutions as long as its content is in no way
RN modified and this statement is not removed. Usage by and for commercial
RN entities requires a license agreement (See http://www.isb-sdb.ch/announce/
RN or send an email to license@sdb.ch).
RN -----
CC CC
CC DR EMBL: M16103; AAA27716.1; -.
CC DR PIR: A26623; A26623.
CC DR HSSP: P02592; 1ET3.
CC DR InterPro: IPR002048; EF-hand.
CC DR Pfam: PF00036; ehand; 3.
CC DR ProDom: PD000012; EF-hand; 1.
CC DR SMART: SM00054; EFh; 2.
CC DR PROSITE: PS00018; EF-HAND; 3.
CC KW Photoprotein; Calcium-binding; Luminescence; Repeat.
CC FT PROPEP 1 7
CC FT CHAIN 8 196
CC FT SITE 47 57 Aequorin 1.
CC FT SITE 62 72 MAY INTERACT WITH THE CHROMOPHORE.
CC FT SITE 107 117 MAY INTERACT WITH THE CHROMOPHORE.
CC FT CA_BIND 31 42 EF-HAND 1 (BY SIMILARITY).

```

```

FT DOMAIN 72 88 ANCESTRAL CALCIUM SITE 2.
FT CA_BIND 124 135 EF-HAND 3 (BY SIMILARITY).
FT CA_BIND 160 171 EF-HAND 4 (BY SIMILARITY).
FT DISULFID 152 159
SQ SEQUENCE 196 AA; 22514 MW; 9AA5B636288A5B8F CRC64;
Query Match 100.0%; Score 27; DB 1; Length 196;
Best Local Similarity 100.0%; Pred No. 54;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGAVP 5
Db 192 GGAVP 196
RESULT 13
AEQ2_AEQVI STANDARD; PRT; 196 AA.
ID AEQ2_AEQVI
AC P02592;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Aequorin 2 precursor.
OS Aequorea victoria (Jellyfish).
OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydrozoa; Leptomedusae;
OC Aequoreidae; Aequorea.
OX NCBI_TaxId=6100;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85216460; PubMed=3858813;
RA Inouye S., Noguchi M., Sakaki Y., Takegi Y., Miyata T., Iwanaga S.,
RA Miyata T., Tsuji F.I.;
RT "Cloning and sequence analysis of cDNA for the luminescent protein
RT aequorin."
RL Proc. Natl. Acad. Sci. U.S.A. 82:3154-3158(1985).
RN [2]
RP SEQUENCE OF 9-185 FROM N.A. (Aequorin 2 AND 3).
RX MEDLINE=87185437; PubMed=2882777;
RA Prasher D.C., McCann R.O., Longiaru M., Cormier M.J.;
RT "Sequence comparisons of complementary DNAs encoding aequorin
RT isotypes."
RL Biochemistry 26:1326-1332(1987).
RN [3]
RP SEQUENCE OF 8-196.
RX MEDLINE=86077721; PubMed=2866797;
RA Charbonneau H., Walsh K.A., McCann R.O., Prendergast F.G.,
RA Cormier M.J., Vanaman T.C.;
RT "Amino acid sequence of the calcium-dependent photoprotein aequorin."
RL Biochemistry 24:6762-6771(1985).
RN [4]
RP MUTAGENESIS.
RN [5]
RP MUTAGENESIS OF PRO-196.
RX MEDLINE=92111761; PubMed=1765170;
RA Nomura M., Inouye S., Ohmura Y., Tsuji F.I.;
RT "A C-terminal proline is required for bioluminescence of the Ca(2+)-
RT binding photoprotein, aequorin."
RL FEBS Lett. 295:63-66(1991).
RN [6]
RP DISULFIDE BOND.
RX MEDLINE=94009705; PubMed=8405461;
RA Ohmura Y., Kurono S., Ohnishi M., Fagan T.F., Tsuji F.I.;
RT "Mass spectrometric evidence for a disulfide bond in aequorin
RT regeneration."
RL FEBS Lett. 332:226-228(1993).
RN [7]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
RX MEDLINE=20289815; PubMed=10830969;
RA Head J.F., Inouye S., Teranishi K., Shimomura O.;

```

```

RT      "The crystal structure of the photoprotein aequorin at 2.3-A
RT      resolution."
CC      -1- FUNCTION: CA(++)-DEPENDENT BIOLUMINESCENCE PHOTOPROTEIN. DISPLAYS
CC      AN EMISSION PEAK AT 470 NM (BLUE LIGHT). TRACE AMOUNTS OF CALCIUM
CC      ION TRIGGER THE INTRAMOLECULAR OXIDATION OF THE CHROMOPHORE,
CC      COELENTERAZINE INTO COELENTERAMIDE AND CO(2) WITH THE
CC      CONCOMITANT EMISSION OF LIGHT.
CC      -1- PFM: THE REDUCTION OF THE DISULFIDE BOND IS NECESSARY TO
CC      REGENERATE Aequorin FROM APOAequorin.
CC      -1- SIMILARITY: BELONGS TO THE EF-HAND CALCIUM-BINDING DOMAINS.
CC      -1- SIMILARITY: CONTAINS 3 EF-HAND CALCIUM-BINDING DOMAINS.
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: L29571; AAA27720.1; -
DR      EMBL: M16104; AAA27717.1; -
DR      EMBL: M16105; AAA27718.1; -
DR      EMBL: M11394; AAA27719.1; -
DR      PIR: A03020; AQTENV.
DR      PDB: 1EJ3; 31-MAY-00.
DR      InterPro: IPR002048; EF-hand.
DR      Pfam: PF000036; efhand; 3.
DR      ProDom: PD000012; EF-hand; 1.
DR      SMART: SM00054; EFh; 2.
DR      PROSITE: PS00018; EF_HAND; 3.
KW      Photoprotein; Calcium-binding; Luminescence; Repeat; 3D-structure.
FT      PROPEP      1      7
FT      CHAIN      1      196
FT      SITE      47      57
FT      SITE      62      72
FT      SITE      107     117
FT      CA_BIND    31      42
FT      DOMAIN     72      88
FT      CA_BIND    124     135
FT      CA_BIND    160     171
FT      DISULFID   152     159
FT      SITE      196     196
FT      VARIANT    70      71
FT      VARIANT    164     164
FT      MUTAGEN    36      36
FT      MUTAGEN    129     129
FT      MUTAGEN    165     165
FT      MUTAGEN    65      65
FT      MUTAGEN    152     152
FT      MUTAGEN    152     152
FT      MUTAGEN    159     159
FT      MUTAGEN    187     187
FT      CONFLICT   37      37
SQ      SEQUENCE    196 AA; 22285 MW; 532DC7A9D29BA80C CRC64;

Query Match      100.0%; Score 27; DB 1; Length 196;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GGAVP 5
DB      192 GGAVP 196

```

```

DE      Mitrocomin precursor.
GN      M17.
OS      Mitrocoma cellularia (Halimastrea mitrocoma).
OC      Eukaryota; Metazoa; Hydrozoa; Hydrozoa; Anthomedusae;
OC      Mitrocomidae; Mitrocoma.
OX      NCBI_TaxID=31874;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=94039781; PubMed=8224198;
RA      Fagan T.F., Ohmura Y., Blinks J.R., Inouye S., Tsuji F.I.;
RT      "Cloning, expression and sequence analysis of cDNA for the Ca(2+)-
RT      binding photoprotein, mitrocomin."
RL      FEBS Lett. 333:301-305(1993).
CC      -1- FUNCTION: CA(++)-DEPENDENT BIOLUMINESCENCE PHOTOPROTEIN. DISPLAYS
CC      AN EMISSION PEAK AT 470 NM (BLUE LIGHT). TRACE AMOUNTS OF CALCIUM
CC      ION TRIGGER THE INTRAMOLECULAR OXIDATION OF THE CHROMOPHORE,
CC      COELENTERAZINE INTO COELENTERAMIDE AND CO(2) WITH THE
CC      CONCOMITANT EMISSION OF LIGHT.
CC      -1- SIMILARITY: BELONGS TO THE EF-HAND CALCIUM-BINDING DOMAINS.
CC      -1- SIMILARITY: CONTAINS 3 EF-HAND CALCIUM-BINDING DOMAINS.
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: L31623; AAA29298.1; -
DR      PIR: S39022; 1EJ3.
DR      HSSP: P02592; 1EJ3.
DR      InterPro: IPR002048; EF-hand.
DR      Pfam: PF00036; efhand; 3.
DR      ProDom: PD000012; EF-hand; 1.
DR      SMART: SM00054; EFh; 2.
DR      PROSITE: PS00018; EF_HAND; 3.
KW      Photoprotein; Calcium-binding; Luminescence; Repeat.
FT      PROPEP      1      8
FT      CHAIN      9      198
FT      CA_BIND    32      43
FT      DOMAIN     73      89
FT      CA_BIND    125     136
FT      CA_BIND    161     172
FT      DISULFID   153     160
SQ      SEQUENCE    198 AA; 22714 MW; 8F6307EF0966F670 CRC64;

Query Match      100.0%; Score 27; DB 1; Length 198;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GGAVP 5
DB      193 GGAVP 197

```

```

RESULT 14
MYR_MITCE
ID MYR_MITCE STANDARD; PRT; 198 AA.
AC P39047;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)

```

```

RESULT 15
OPBB_BACSU
ID OPBB_BACSU STANDARD; PRT; 217 AA.
AC Q45461; Q34670;
DT 01-NOV-1997 (Rel. 35, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Choline transport system permease protein opubb.
GN OPUBB OR PROM.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 6633 / LHA5;
RX MEDLINE=96074318; PubMed=7592481;
RA Llin Y., Hansen J.N.;

```

RT "Characterization of a chimeric *proU* operon in a subtilin-producing
 RT mutant of *Bacillus subtilis* 168." ;
 RL J. Bacteriol. 177:6874-6880(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168 / JH642;
 RX MEDLINE=99232519; PubMed=10216873;
 RA Kappes R.M., Kempf B., Kneip S., Boch J., Gade J., Meier-Wagner J.,
 RA Bremer E.;
 RT "Two evolutionarily closely related ABC transporters mediate the
 RT uptake of choline for synthesis of the osmoprotectant glycine betaine
 RT in *Bacillus subtilis*." ;
 RL Mol. Microbiol. 33:203-216(1999).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Bertolo M.G., Bessieres P., Bolotin A., Borchert S.,
 RA Borries R., Bourcier L., Brans A., Braun M., Brignell S.C., Bron S.,
 RA Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
 RA Dentzot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
 RA Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holstappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
 RA Joris B., Karamata D., Kasahara Y., Klier-Blanchard M., Klein C.,
 RA Kobayashi Y., Koelter P., Koningsstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogilwa A., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portelle D., Rapoport G., Prescott A.M.,
 RA Prescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scofield F.,
 RA Sekiguchi J., Sekowska A., Seror S.J., Serro P., Shin B.S., Soldo B.,
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Takakoshi A., Tanaka T., Terpstra P., Tognoni A.,
 RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
 RA Viari A., Wambuit R., Wedler E., Wedler H., Wellenreger T.,
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasunoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
 RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*
 RT *subtilis*." ;
 RL Nature 390:249-256(1997).
 CC -1- FUNCTION: INVOLVED IN A HIGH AFFINITY MULTICOMPONENT BINDING-
 CC PROTEIN-DEPENDENT TRANSPORT SYSTEM FOR CHOLINE; PROBABLY
 CC RESPONSIBLE FOR THE TRANSLLOCATION OF THE SUBSTRATE ACROSS THE
 CC MEMBRANE.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO THE BINDING-PROTEIN-DEPENDENT TRANSPORT
 CC SYSTEM PERMEASE FAMILY. CYSTM SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: U38418; AB01533.1; -;
 DR EMBL: AF008930; AAC14357.1; -;
 DR EMBL: Z99121; CAB15377.1; -;
 DR Subtilist; BG12634; opub.
 DR InterPro: IPR000515; BPD_transp.
 DR Pfam: PF00528; BPD_transp.1.
 DR PROSITE: PS00402; BPD_TRANSNIP_INN_MEMBER.1.
 KW Transport; Amino-acid transport; Transmembrane; Complete proteome.
 FT TRANSMEM 23 43
 FT POTENTIAL.

FT	TRANSMEM	72	92	POTENTIAL.
FT	TRANSMEM	128	148	POTENTIAL.
FT	TRANSMEM	150	170	POTENTIAL.
FT	TRANSMEM	180	200	POTENTIAL.
FT	VARIANT	5	5	POTENTIAL.
FT	VARIANT	130	131	EL -> DV (IN STRAIN LH45).
FT	VARIANT	201	201	A -> T (IN STRAIN LH45).
FT	VARIANT	216	216	L -> V (IN STRAIN LH45).
FT	CONFLECT	67	72	LA1AF -> PDSRH (IN REF. 1).
SQ	SEQUENCE	217 AA;	23163 MW;	A7915A2D22ABA58E CRC64;

Query Match 100.0%; Score 27; DB 1; Length 217;
 Best Local Similarity 100.0%; Pred. No. 60;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAVP 5
 DB 182 GGAVP 186

Search completed: March 17, 2003, 12:07:19
 Job time : 9 secs

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 17, 2003, 12:05:48 ; Search time 25 Seconds
(without alignments)
41.209 Million cell updates/sec

Title: SEQ25-AV
Perfect score: 27
Sequence: 1 ggapv 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

SPTRMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_protein:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriophage:*
17: sp_archaeal:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	27	100.0	53	6 Q9N098	Q9N098 macaca fasc
2	27	100.0	54	2 Q9S4W8	Q9S4W8 escherichia
3	27	100.0	62	10 Q8S7J8	Q8S7J8 oryza sativ
4	27	100.0	63	10 Q9XEX1	Q9XEX1 dendrobium
5	27	100.0	83	12 Q69211	Q69211 human cytom
6	27	100.0	91	10 Q9LED1	Q9LED1 prunus pers
7	27	100.0	92	2 Q54077	Q54077 saccharopol
8	27	100.0	100	5 Q26874	Q26874 taenia cras
9	27	100.0	100	5 Q26876	Q26876 taenia cras
10	27	100.0	100	9 Q9T220	Q9T220 bacterioph
11	27	100.0	100	9 Q37983	Q37983 bacterioph
12	27	100.0	103	16 Q9L063	Q9L063 streptomyce
13	27	100.0	105	16 Q8Y0X7	Q8Y0X7 raietonia s
14	27	100.0	107	17 Q9YD12	Q9YD12 aeropyrum p
15	27	100.0	112	10 Q8S2S8	Q8S2S8 thelungtel
16	27	100.0	116	2 Q68056	Q68056 rhodobacter

17	27	100.0	117	10 Q24440	Q24440 phaseolus v
18	27	100.0	117	10 Q8W539	Q8W539 rietama raet
19	27	100.0	117	10 Q8VX12	Q8VX12 fragaria an
20	27	100.0	128	10 Q9S0Q8	Q9S0Q8 arabidopsis
21	27	100.0	132	11 Q9D5R1	Q9D5R1 mus musculu
22	27	100.0	132	16 Q92JW2	Q92JW2 rhizobium m
23	27	100.0	134	2 Q50085	Q50085 mycobacteri
24	27	100.0	134	2 Q931L24	Q931L24 mycobacteri
25	27	100.0	145	5 Q17890	Q17890 caenorhabd1
26	27	100.0	147	2 Q9K5H6	Q9K5H6 bordetella
27	27	100.0	150	2 Q9K5H8	Q9K5H8 bordetella
28	27	100.0	150	2 Q9K5H4	Q9K5H4 bordetella
29	27	100.0	154	2 Q9K5H0	Q9K5H0 bordetella
30	27	100.0	154	5 Q9VD12	Q9VD12 drosophila
31	27	100.0	155	2 Q9K5H2	Q9K5H2 bordetella
32	27	100.0	158	2 Q9K5G6	Q9K5G6 bordetella
33	27	100.0	159	2 Q9K5G8	Q9K5G8 bordetella
34	27	100.0	159	2 Q9ALP4	Q9ALP4 bordetella
35	27	100.0	160	2 Q9K5T0	Q9K5T0 bordetella
36	27	100.0	160	2 Q9K5G3	Q9K5G3 bordetella
37	27	100.0	161	2 Q9K5G2	Q9K5G2 bordetella
38	27	100.0	167	2 Q937U5	Q937U5 bordetella
39	27	100.0	173	2 Q9ALQ5	Q9ALQ5 bordetella
40	27	100.0	175	17 Q26596	Q26596 bordetella
41	27	100.0	177	2 Q9ALQ3	Q9ALQ3 methanobact
42	27	100.0	178	2 Q93ST1	Q93ST1 bordetella
43	27	100.0	179	2 Q9AHP1	Q9AHP1 bordetella
44	27	100.0	180	6 Q95LA0	Q95LA0 bos taurus
45	27	100.0	181	2 Q9ALQ1	Q9ALQ1 bordetella

ALIGNMENTS

RESULT 1
ID Q9N098 PRELIMINARY; PRT; 53 AA.
AC Q9N098;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE Unnamed protein product.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RA Oseada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Teroo K.,
RA Suzuki Y., Sugano S., Hashimoto K.;
RT "Isolation of full-length cDNA clones from macaque brain cDNA
RT libraries".
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB046032; BAB01614.1;
SQ SEQUENCE 53 AA; 5387 MW; E5E48B2EE9248A6E CRC64;

Query Match Score 27; DB 6; Length 53;
Best Local Similarity 100.0%; Pred. No. 11e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNAV 5
Db 42 GNAV 46

RESULT 2
ID Q9S4W8 PRELIMINARY; PRT; 54 AA.
AC Q9S4W8;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)

DE Hypothetical 5.3 kDa protein.
 OS Escherichia coli.
 OG Plasmid F.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9296678; PubMed=10366527;
 RA Manwaring N.P., Skurray R.A., Firth N.;
 RT "Nucleotide sequence of the F plasmid leading region."
 RL Plasmid 41:219-225(1999).
 DR EMBL: AF106329; AAD47182.1; -
 KW Hypothetical protein; Plasmid.
 SQ SEQUENCE 54 AA; 5304 MW; A66D08C23346A415 CRC64;

Query Match 100.0%; Score 27; DB 10; Length 54;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGAVP 5
 DB 48 GGAVP 52

RESULT 3
 ID 0857J8 PRELIMINARY; PRT; 62 AA.
 AC 0857J8;
 DT 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Hypothetical 6.1 kDa protein.
 GN OSUBA0095C06.10.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzeae; Oryza.
 OX NCBI_TaxID=45530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NIPPONBARE;
 RA Du H., Minx P., de la Bastide M., Nascimento L., Spiegel L.,
 RA Preston R., Kirchoff K., Kuit K., Baker J., Vili M.D., Zuberav T.,
 RA Santos L., Cunniss D.M., Rodriguez S., Miller B., Balija V., Shah R.,
 RA Baheti A., King L., O'Shaughnessy A., Palmer L., Dedhia N.,
 RA McCombe W.R.;
 RT "Genomic sequence for Oryza sativa, Nipponbare strain, clone
 RT OSUBA0095C06, from Chromosome 10, complete sequence."
 RL Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC079935; AAM08538.1; -
 KW Hypothetical protein.
 SQ SEQUENCE 62 AA; 6120 MW; 59DB8339A3D923D CRC64;

Query Match 100.0%; Score 27; DB 10; Length 62;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGAVP 5
 DB 22 GGAVP 26

RESULT 4
 ID 09XEX1 PRELIMINARY; PRT; 63 AA.
 AC 09XEX1;
 DT 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE Putative alternative oxidase (Fragment).
 GN OTG9.
 OS Dendrobium grex Madame Thong-In.

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Orchidaceae;
 OC Epidendroideae; higher Epidendroideae; Dendrobieae; Dendrobilinae;
 OC Dendrobium.
 OX NCBI_TaxID=84618;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yu H., Goh C.J.;
 RT "Characterization of genes differentially expressed during orchid
 RT floral transition."
 RL Submitted (Nov-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF107589; AAD20817.1; -
 FT NON_TER 1 1
 FT NON_TER 63 63
 KW Hypothetical protein.
 SQ SEQUENCE 63 AA; 6682 MW; BA02489B4FB6B35 CRC64;

Query Match 100.0%; Score 27; DB 10; Length 63;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGAVP 5
 DB 49 GGAVP 53

RESULT 5
 ID 069211 PRELIMINARY; PRT; 83 AA.
 AC 069211;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-NOV-1996 (TREMBlrel. 08, Last annotation update)
 DE Hypothetical 9.5 kDa protein.
 OS Human cytomegalovirus.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Betaherpesvirinae; Cytomegalovirus.
 OX NCBI_TaxID=10359;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88289786; PubMed=2840673;
 RA Razaque A., Jahan N., Mcweeney D., Jariwalla R.J., Jones C.,
 RA Brady J., Rosenthal L.J.;
 RT "Localization and DNA sequence analysis of the transforming domain
 RT (mrit) of human cytomegalovirus."
 RL Proc. Natl. Acad. Sci. U.S.A. 85:5709-5713(1988).
 DR EMBL: J03822; AAA6544.1; -
 KW Hypothetical protein.
 SQ SEQUENCE 83 AA; 9455 MW; 4F94A00F976B57D3 CRC64;

Query Match 100.0%; Score 27; DB 12; Length 83;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGAVP 5
 DB 55 GGAVP 59

RESULT 6
 ID 09LED1 PRELIMINARY; PRT; 91 AA.
 AC 09LED1;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Pru p 1 (Fragment).
 GN PRUPL.
 OS Prunus persica (Peach).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Eufrosids I; Rosales; Rosaceae; Amygaloideae; Prunus.
 OX NCBI_TaxID=3760;
 RN [1]

RP SEQUENCE FROM N.A.
 RA Diaz-Berles A., Garcia-Casado G., Sanchez-Monge R., Barber D.,
 RA Salcedo G.,
 RT "Cloning of a cDNA encoding a major allergen and lipid transfer
 RT protein, from *Persica vulgaris* Mill.,"
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ277163; CAB96876.2; -
 DR HSSP: P19656; 1MZM
 DR InterPro: IPR003612; AAI.
 DR InterPro: IPR000528; Plant_LTP.
 DR InterPro: IPR001768; Try/amyL_inhbr.
 DR Pfam: PF00234; Tryp_alpha_amy1; 1.
 DR PRINTS: PR00382; LIPIDTRANSF.
 DR SMART: SM00499; AAI; 1.
 DR PROSITE: PS00597; PLANT_LTP; 1.
 FT NON_TER 1 1
 SQ SEQUENCE 91 AA; 9143 MW; C8944885481AACCE CRC64;

Query Match 100.0%; Score 27; DB 10; Length 91;
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGAVP 5
 DB 20 GGAVP 24

RESULT 7

ID 054077 PRELIMINARY; PRT; 92 AA.
 AC 054077;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 07-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE 0-ff2.
 GN ORF2.
 OS Saccharopolyspora erythraea (Streptomyces erythraeus).
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Pseudonocardineae; Pseudonocardiaceae;
 OC Saccharopolyspora.
 OX NCBI_TaxID=1836;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94211208; PubMed=8159169;
 RA Brown D.P., Idler K.B., Backer D.M., Donadio S., Katz L.;
 RT "Characterization of the genes and attachment sites for site-specific
 RT integration of plasmid pSEL101 in *Saccharopolyspora erythraea* and
 RT *Streptomyces lividans*,"
 RL Mol. Gen. Genet. 242:185-193(1994).
 DR EMBL: L11597; AAA26478.1; -
 SQ SEQUENCE 92 AA; 9954 MW; FEB07A8769E3AB8F CRC64;

Query Match 100.0%; Score 27; DB 2; Length 92;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGAVP 5
 DB 17 GGAVP 21

RESULT 8

ID 026874 PRELIMINARY; PRT; 100 AA.
 AC 026874;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Protective recombinant antigen (Fragment).
 GN KETC7.
 OS Taenia crassiceps.
 OC Eukaryota; Metazoa; Platyhelminthes; Cestoda; Eucestoda;
 OC Cyclophyllidae; Taeniidae; Taenia.

OX NCBI_TaxID=6207;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Manoucharian K.G., Rosas G., Hernandez M.R., Fragoso G., Aluja A.,
 RA Villalobos N., Rodarte L.F., Scuitto E.;
 RT "Cysticercosis: Identification and cloning of protective recombinant
 RT antigens,"
 RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U30899; AAA74982.1; -
 FT NON_TER 1 1
 SQ SEQUENCE 100 AA; 10636 MW; F7F379F871CDBD2B CRC64;

Query Match 100.0%; Score 27; DB 5; Length 100;
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGAVP 5
 DB 31 GGAVP 35

RESULT 9

ID 026876 PRELIMINARY; PRT; 100 AA.
 AC 026876;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
 DE Protective recombinant antigen (Fragment).
 OS Taenia crassiceps.
 OC Eukaryota; Metazoa; Platyhelminthes; Cestoda; Eucestoda;
 OC Cyclophyllidae; Taeniidae; Taenia.
 OX NCBI_TaxID=6207;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ORF;
 RX MEDLINE=96185764; PubMed=8604092;
 RA Manoucharian K., Rosas G., Hernandez M., Fragoso G., Aluja A.,
 RA Villalobos N., Rodarte L.F., Scuitto E.;
 RT "Cysticercosis: Identification and cloning of protective recombinant
 RT antigens,"
 RL J. Parasitol. 82:250-254(1996).
 DR EMBL: U31524; AAB02180.1; -
 FT NON_TER 1 1
 SQ SEQUENCE 100 AA; 10594 MW; B6F365886DA1CD37 CRC64;

Query Match 100.0%; Score 27; DB 5; Length 100;
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGAVP 5
 DB 31 GGAVP 35

RESULT 10

ID 09T220 PRELIMINARY; PRT; 100 AA.
 AC 09T220;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE GP27.
 GN Bacteriophage phi-C31.
 OS Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
 OC Lambda-like viruses.
 OX NCBI_TaxID=10719;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-NORWICH STOCK;
 RA Smith M.C.M.;
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.

```

RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-NORMWICH STOCK;
RX MEDLINE-99162580; PubMed-10051617;
RA Hendrix R.W., Smith M.C.M., Burns N., Ford M.E., Hatfull G.F.;
RT "All the world's a phage.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:2192-2197(1999).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-NORMWICH STOCK;
RX MEDLINE-99238410; PubMed-10219087;
RA Smith M.C.M., Burns N., Wilson R.N., Gregory M.A.;
RT "The complete genome sequence of the Streptomyces temperate phage C31:
evolutionary relationships to other viruses.";
RL Nucleic Acids Res. 27:2145-2155(1999).
DR EMBL; AJ006589; CAA07151.1; -.
SQ SEQUENCE 100 AA; 11558 MW; AE24050FE7910B5 CRC64;

Query Match
Best Local Similarity 100.0%; Score 27; DB 9; Length 100;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAVP 5
DB 86 GGAVP 90

RESULT 11
Q37983 PRELIMINARY; PRT; 100 AA.
ID 037983;
AC 037983;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE ORF2 protein.
OS Bacteriophage phi-C31.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
OC Lambda-like viruses.
OX NCBI_TaxID=10719;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-92020203; PubMed-1656389;
RA Rausch H., Lehmann M.;
RT "Structural analysis of the actinophage phi C31 attachment site.";
RL Nucleic Acids Res. 19:5187-5189(1991).
DR EMBL; X57036; CAA40352.1; -.
SQ SEQUENCE 100 AA; 11548 MW; FD273563FDA10B7 CRC64;

Query Match
Best Local Similarity 100.0%; Score 27; DB 9; Length 100;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAVP 5
DB 86 GGAVP 90

RESULT 12
Q9L063 PRELIMINARY; PRT; 103 AA.
ID 09L063;
AC 09L063;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)
DE Hypothetical protein SC02791.
GN SC02791 OR SC0105.22C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2) / M145;

```

```

RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Gobie A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kleser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AL163641; CAB87228.1; -.
KW Hypothetical protein.
SQ SEQUENCE 103 AA; 10639 MW; 9316F1F38C1BD492 CRC64;

Query Match
Best Local Similarity 100.0%; Score 27; DB 16; Length 103;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAVP 5
DB 62 GGAVP 66

RESULT 13
Q8Y0X7 PRELIMINARY; PRT; 105 AA.
ID 08Y0X7;
AC 08Y0X7;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, last annotation update)
DE Putative COMEA-related signal peptide protein.
GN RSC0916 OR RS04499.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
OC Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-GMT1000;
RX MEDLINE-21681879; PubMed-11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Arlat M., Billault A., Broctier P., Camus J.C., Cattolico L.,
RA Chandler M., Choisme N., Claudel-Renard C., Cunac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Sautin W., Schlex T.,
RA Sigulier P., Thebaud P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502(2002).
DR EMBL; AL646061; CAD14618.1; -.
KW Complete proteome.
SQ SEQUENCE 105 AA; 10968 MW; FCD8031E7F4B6AFD CRC64;

Query Match
Best Local Similarity 100.0%; Score 27; DB 16; Length 105;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAVP 5
DB 88 GGAVP 92

RESULT 14
Q9YD12 PRELIMINARY; PRT; 107 AA.
ID 09YD12;
AC 09YD12;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, last annotation update)
DE Hypothetical protein APE1100.
GN APE1100.
OS Aeropyrum pernix.

```



```

OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcaceae;
OC Desulfurococcaceae; Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-K1;
RX MEDLINE-99310339; PubMed-10382966;
RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Halkawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankal A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai S., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudo Y.,
RA Yamazaki J., Kishida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT *Complete genome sequence of an aerobic hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix K1.;
RL DNA Res. 6:83-101(1999).
DR EMBL; AF000060; BAA80085.1;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 107 AA; 11451 MW; 8296C1B7152AAE64 CRC64;

Query Match          100.0%; Score 27; DB 17; Length 107;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAVP 5
   |||||
Db 72 GGAVP 76

RESULT 15
Q8S2S8      PRELIMINARY;      PRT;      112 AA.
ID Q8S2S8;
AC Q8S2S8;
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Lipid transfer protein 4-like protein.
OS Theilungella halophila.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eunostids II; Brassicales; Brassicaceae; Theilungella.
OX NCBI_TaxID=96038;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhang C., Gong Z., Kim C., Zhu J.-K.;
RT *cDNA clones from Theilungella halophila.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF499715; AAM19702.1;
SQ SEQUENCE 112 AA; 11213 MW; BD7CEC88191B4AB4 CRC64;

Query Match          100.0%; Score 27; DB 10; Length 112;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAVP 5
   |||||
Db 44 GGAVP 48

```

Search completed: March 17, 2003, 12:10:24
 Job time : 28 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 17, 2003, 12:06:59 ; Search time 12 Seconds
(without alignments)
12.260 Million cell updates/sec

Title: SEQ25-AV
Perfect score: 27
Sequence: 1 ggapv 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/3A_COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/3B_COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/4A_COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/4B_COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCUTS_COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfillseq.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	27	100.0	38	2	US-08-448-481-1 Sequence 1, Appl1
2	27	100.0	167	4	US-09-199-637A-361 Sequence 361, App
3	27	100.0	189	1	US-07-982-650D-2 Sequence 2, Appl1
4	27	100.0	189	1	US-08-331-379-2 Sequence 1, Appl1
5	27	100.0	196	1	US-07-982-650D-1 Sequence 3, Appl1
6	27	100.0	196	1	US-07-982-650D-3 Sequence 4, Appl1
7	27	100.0	196	1	US-07-982-650D-4 Sequence 5, Appl1
8	27	100.0	196	1	US-07-982-650D-5 Sequence 6, Appl1
9	27	100.0	196	1	US-08-331-379-1 Sequence 1, Appl1
10	27	100.0	196	1	US-08-331-379-3 Sequence 3, Appl1
11	27	100.0	196	1	US-08-331-379-5 Sequence 5, Appl1
12	27	100.0	196	1	US-08-331-379-6 Sequence 6, Appl1
13	27	100.0	283	1	US-08-583-672-2 Sequence 2, Appl1
14	27	100.0	284	2	US-08-320-148B-2 Sequence 2, Appl1
15	27	100.0	284	3	US-08-589-028-6 Sequence 6, Appl1
16	27	100.0	284	3	US-08-784-582-6 Sequence 6, Appl1
17	27	100.0	284	4	US-08-785-271-6 Sequence 6, Appl1
18	27	100.0	284	4	US-09-031-898-2 Sequence 2, Appl1
19	27	100.0	340	4	US-09-214-631-3 Sequence 3, Appl1
20	27	100.0	349	4	US-09-343-011B-1 Sequence 1, Appl1
21	27	100.0	415	4	US-09-180-109A-9 Sequence 9, Appl1
22	27	100.0	415	4	US-09-180-109A-12 Sequence 12, Appl1
23	27	100.0	417	4	US-09-180-109A-10 Sequence 10, Appl1
24	27	100.0	417	4	US-09-180-109A-14 Sequence 14, Appl1
25	27	100.0	449	4	US-08-697-954-4 Sequence 6, Appl1
26	27	100.0	524	1	US-08-336-257A-6 Patent No. 5386025
27	27	100.0	524	6	5386025-4

28	27	100.0	539	4	US-09-036-987A-11 Sequence 11, Appl1
29	27	100.0	539	4	US-09-370-700-11 Sequence 11, Appl1
30	27	100.0	543	2	US-08-469-412A-7 Sequence 7, Appl1
31	27	100.0	543	4	US-09-021-715-7 Sequence 7, Appl1
32	27	100.0	548	2	US-08-469-412A-2 Sequence 2, Appl1
33	27	100.0	548	4	US-09-021-715-2 Sequence 2, Appl1
34	27	100.0	554	4	US-08-895-590-9 Sequence 5, Appl1
35	27	100.0	910	4	US-08-460-269C-2 Sequence 4, Appl1
36	27	100.0	911	4	US-08-460-269C-4 Sequence 4, Appl1
37	27	100.0	922	4	US-08-460-269C-6 Sequence 6, Appl1
38	27	100.0	1276	1	US-08-222-616-24 Sequence 24, Appl1
39	27	100.0	1276	4	US-08-446-648-24 Sequence 24, Appl1
40	27	100.0	1276	5	PCT-US95-04228-24 Sequence 24, Appl1
41	27	100.0	1395	4	US-09-340-245A-15 Sequence 15, Appl1
42	27	100.0	2512	2	US-08-801-263A-9 Sequence 9, Appl1
43	27	100.0	2512	3	US-09-102-248-9 Sequence 9, Appl1
44	27	100.0	3724	2	US-08-804-227C-10 Sequence 10, Appl1
45	27	100.0	3724	2	US-08-804-198-4 Sequence 4, Appl1

ALIGNMENTS

RESULT 1
US-08-448-481-1
Sequence 1, Application US/08448481
Patent No. 5914270
GENERAL INFORMATION:
APPLICANT: Coutos-Thévenot, Pierre M.L.
APPLICANT: Jouenne, Thierry G.
APPLICANT: Maes, Olivier C.A.
APPLICANT: Deloire, Alain J.
APPLICANT: Boulay, Michel P.H.
APPLICANT: Guern, Jean R.D.
TITLE OF INVENTION: Method for Promoting the Differentiation
of Plant Cells in Culture
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSER: Pollock, Vande Sande & Priddy
STREET: 1990 W Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20036-3425
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/448,481
FILING DATE: 08-AUG-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR93/01239
FILING DATE: 14-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 92 15044
FILING DATE: 14-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Franklin, Eric J.
REGISTRATION NUMBER: 37,134
REFERENCE/DOCKET NUMBER: 0061/00065
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 331-7111
TELEFAX: (202) 293-6229
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide

US-08-448-481-1

Query Match 100.0%; Score 27; DB 2; Length 38;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGAVP 5
|||||
DB 21 GGAVP 25

RESULT 2

US-09-199-637A-361
; Sequence 361, Application US/09199637A
; Patent No. 6355411
; GENERAL INFORMATION:
; APPLICANT: Ausubel, Frederick
; APPLICANT: Goodman, Howard M.
; APPLICANT: Rahne, Laurence G.
; APPLICANT: Mahajan-Miklos, Shalina
; APPLICANT: Cao, Hui
; APPLICANT: Tan, Man-Wah
; APPLICANT: Drenkard, Eliana
; APPLICANT: Tsongalis, John
; TITLE OF INVENTION: VIRUENCE-ASSOCIATED NUCLEIC ACID
; TITLE OF INVENTION: SEQUENCES AND USES THEREOF
; FILE REFERENCE: 00786/361002
; CURRENT APPLICATION NUMBER: US/09/199,637A
; PRIOR FILING DATE: 1998-11-25
; PRIOR APPLICATION NUMBER: 60/066,517
; NUMBER OF SEQ ID NOS: 437
; SOFTWARE: PastSeq for Windows Version 4.0
; SEQ ID NO 361
; LENGTH: 167
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-199-637A-361

Query Match 100.0%; Score 27; DB 4; Length 167;
Best Local Similarity 100.0%; Pred. No. 1,8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGAVP 5
|||||
DB 39 GGAVP 43

RESULT 3

US-07-982-650D-2
; Sequence 2, Application US/07982650D
; Patent No. 5360728
; GENERAL INFORMATION:
; APPLICANT: Prasher, Douglas
; TITLE OF INVENTION: Modified Aequorin Having Increased
; TITLE OF INVENTION: Bioluminescent Activity
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30303-1769
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch
; OPERATING SYSTEM: 7.0
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/982,650D
; FILING DATE: December 1, 1992
; CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: No. 5360728e
FILING DATE: No. 5360728e
ATTORNEY/AGENT INFORMATION:
NAME: Larry W. Stults, Ph.D.
REGISTRATION NUMBER: 34,025
REFERENCE/DOCKET NUMBER: 04840-0070
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-818-3700
TELEFAX: 404-818-3799
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 189 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Recombinant Aequorin
OTHER INFORMATION: Bioluminescent protein
PUBLICATION INFORMATION:
AUTHORS: Chaboudneau et al.
TITLE: Amino acid sequence of the calcium-dependent photoprotein aequorin
JOURNAL: Am. Chem. Soc.
VOLUME: 24
ISSUE: 24
PAGES: 6762-6771
DATE: 1985
RELEVANT RESIDUES IN SEQ ID NO: 2: Points of microheterogeneity where specif

Query Match 100.0%; Score 27; DB 1; Length 189;
Best Local Similarity 100.0%; Pred. No. 2,1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGAVP 5
|||||
DB 185 GGAVP 189

RESULT 4

US-08-331-379-2
; Sequence 2, Application US/08331379
; Patent No. 5541309
; GENERAL INFORMATION:
; APPLICANT: Prasher, Douglas
; TITLE OF INVENTION: Modified Aequorin Having Increased
; TITLE OF INVENTION: Bioluminescent Activity
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30303-1769
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch
; OPERATING SYSTEM: 7.0
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/331,379
; FILING DATE: October 28, 1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: No. 5541309e
; FILING DATE: No. 5541309e
; ATTORNEY/AGENT INFORMATION:
NAME: Larry W. Stults, Ph.D.
REGISTRATION NUMBER: 34,025
REFERENCE/DOCKET NUMBER: 04840-0071
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-818-3700

TELEFAX: 404-818-3799
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 189 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: NAME/KEY: Recombinant Aequorin
: OTHER INFORMATION: Bioluminescent protein
: PUBLICATION INFORMATION:
: AUTHORS: Charbonneau et al.
: TITLE: Amino acid sequence of the calcium-dependent photoprotein aequorin
: JOURNAL: Am. Chem. Soc.
: VOLUME: 24
: ISSUE: 24
: PAGES: 6762-6771
: DATE: 1985
: RELEVANT RESIDUES IN SEQ ID NO: 2: Points of microheterogeneity where specific
US-08-331-379-2

Query Match 100.0%; Score 27; DB 1; Length 189;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGAVP 5
|||||

Db 185 GGAVP 189

RESULT 5
US-07-982-650D-1
: Sequence 1, Application US/07982650D
: Patent No. 5360728
: GENERAL INFORMATION:
: APPLICANT: Prasher, Douglas
: TITLE OF INVENTION: Modified Aequorin Having Increased
: TITLE OF INVENTION: Bioluminescent Activity
: NUMBER OF SEQUENCES: 10
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Jones & Askew
: STREET: 191 Peachtree Street, 37th Floor
: CITY: Atlanta
: STATE: Georgia
: COUNTRY: USA
: ZIP: 30303-1769
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette, 3.50 Inch
: COMPUTER: Macintosh
: OPERATING SYSTEM: 7.0
: SOFTWARE: Microsoft Word
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/982, 650D
: FILING DATE: December 1, 1992
: CLASSIFICATION: 530
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: No. 5360728e
: FILING DATE: No. 5360728e
: ATTORNEY/AGENT INFORMATION:
: NAME: Larry W. Stults, Ph.D.
: REGISTRATION NUMBER: 34,025
: REFERENCE/DOCKET NUMBER: 04840-0070
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 404-818-3700
: TELEFAX: 404-818-3799
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 196 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: NAME/KEY: Recombinant Aequorin

OTHER INFORMATION: Bioluminescent protein
: PUBLICATION INFORMATION:
: AUTHORS: Prasher et al.
: TITLE: Sequence Comparisons of Complementary DNAs Encoding
: TITLE: Aequorin Isozymes
: JOURNAL: Biochemistry
: VOLUME: 26
: PAGES: 1326-1332
: DATE: 1987
: RELEVANT RESIDUES IN SEQ ID NO: 1: Points of microheterogeneity where specific
US-07-982-650D-1

Query Match 100.0%; Score 27; DB 1; Length 196;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGAVP 5
|||||

Db 192 GGAVP 196

RESULT 6
US-07-982-650D-3
: Sequence 3, Application US/07982650D
: Patent No. 5360728
: GENERAL INFORMATION:
: APPLICANT: Prasher, Douglas
: TITLE OF INVENTION: Modified Aequorin Having Increased
: TITLE OF INVENTION: Bioluminescent Activity
: NUMBER OF SEQUENCES: 10
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Jones & Askew
: STREET: 191 Peachtree Street, 37th Floor
: CITY: Atlanta
: STATE: Georgia
: COUNTRY: USA
: ZIP: 30303-1769
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette, 3.50 Inch
: COMPUTER: Macintosh
: OPERATING SYSTEM: 7.0
: SOFTWARE: Microsoft Word
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/982, 650D
: FILING DATE: December 1, 1992
: CLASSIFICATION: 530
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: No. 5360728e
: FILING DATE: No. 5360728e
: ATTORNEY/AGENT INFORMATION:
: NAME: Larry W. Stults, Ph.D.
: REGISTRATION NUMBER: 34,025
: REFERENCE/DOCKET NUMBER: 04840-0070
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 404-818-3700
: TELEFAX: 404-818-3799
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 196 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: NAME/KEY: Recombinant site-directed aequorin mutant
: OTHER INFORMATION: Site-directed mutant having increased bioluminescent activity
: PUBLICATION INFORMATION:
: RELEVANT RESIDUES IN SEQ ID NO: 3: Asp 124 changed to Ser
US-07-982-650D-3

Query Match 100.0%; Score 27; DB 1; Length 196;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGAVP 5
DB 192 GGAVP 196

RESULT 7

US-07-982-650D-4
Sequence 4, Application US/07982650D
Patent No. 5360728
GENERAL INFORMATION:
APPLICANT: Prasher, Douglas
TITLE OF INVENTION: Modified Aequorin Having Increased
Bioluminescent Activity
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Jones & Askew
STREET: 191 Peachtree Street, 37th Floor
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30303-1769
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch
COMPUTER: Macintosh
OPERATING SYSTEM: 7.0
SOFTWARE: Microsoft Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/982,650D
FILING DATE: December 1, 1992
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: No. 5360728e
FILING DATE: No. 5360728e
ATTORNEY/AGENT INFORMATION:
NAME: Larry W. Stults, Ph.D.
REGISTRATION NUMBER: 34,025
REFERENCE/DOCKET NUMBER: 04840-0070
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-818-3700
TELEFAX: 404-818-3799
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 196 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Recombinant site-directed aequorin mutant
OTHER INFORMATION: Site-directed aequorin mutant having increased bioluminescent
PUBLICATION INFORMATION:
RELEVANT RESIDUES IN SEQ ID NO: 4: Glu 135 changed to Ser
US-07-982-650D-4

Query Match 100.0%; Score 27; DB 1; Length 196;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGAVP 5
DB 192 GGAVP 196

RESULT 8

US-07-982-650D-5
Sequence 5, Application US/07982650D
Patent No. 5360728
GENERAL INFORMATION:
APPLICANT: Prasher, Douglas
TITLE OF INVENTION: Modified Aequorin Having Increased
Bioluminescent Activity
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Jones & Askew

STREET: 191 Peachtree Street, 37th Floor
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30303-1769
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch
COMPUTER: Macintosh
OPERATING SYSTEM: 7.0
SOFTWARE: Microsoft Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/982,650D
FILING DATE: December 1, 1992
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: No. 5360728e
FILING DATE: No. 5360728e
ATTORNEY/AGENT INFORMATION:
NAME: Larry W. Stults, Ph.D.
REGISTRATION NUMBER: 34,025
REFERENCE/DOCKET NUMBER: 04840-0070
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-818-3700
TELEFAX: 404-818-3799
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 196 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Recombinant site-directed aequorin mutant
OTHER INFORMATION: Site-directed mutant having increased bioluminescent activity
PUBLICATION INFORMATION:
RELEVANT RESIDUES IN SEQ ID NO: 5: Gly 129 changed to Ala
US-07-982-650D-5

Query Match 100.0%; Score 27; DB 1; Length 196;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGAVP 5
DB 192 GGAVP 196

RESULT 9

US-08-331-379-1
Sequence 1, Application US/08331379
Patent No. 5541309
GENERAL INFORMATION:
APPLICANT: Prasher, Douglas
TITLE OF INVENTION: Modified Aequorin Having Increased
Bioluminescent Activity
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Jones & Askew
STREET: 191 Peachtree Street, 37th Floor
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30303-1769
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch
COMPUTER: Macintosh
OPERATING SYSTEM: 7.0
SOFTWARE: Microsoft Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/331,379
FILING DATE: October 28, 1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: No. 5541309e

FILING DATE: No. 5541309
ATTORNEY/AGENT INFORMATION:
NAME: Larry W. Stults, Ph.D.
REGISTRATION NUMBER: 34,025
REFERENCE/DOCKET NUMBER: 04840-0071
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-818-3799
TELEFAX: 404-818-3799
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 196 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Recombinant Aequorin
OTHER INFORMATION: Bioluminescent protein
PUBLICATION INFORMATION:
AUTHORS: Prasher et al.
TITLE: Sequence Comparisons of Complementary DNAs Encoding
Patent No. 5541309
TITLE: Aequorin Isotypes
JOURNAL: Biochemistry
VOLUME: 26
PAGES: 1326-1332
DATE: 1987
RELEVANT RESIDUES IN SEQ ID NO: 1: Points of microheterogeneity where specific
US-08-331-379-1
Query Match 100.0%; Score 27; DB 1; Length 196;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGAVP 5
|||||
Db 192 GGAVP 196
RESULT 10
US-08-331-379-3
Sequence 3, Application US/08331379
Patent No. 5541309
GENERAL INFORMATION:
APPLICANT: Prasher, Douglas
TITLE OF INVENTION: Modified Aequorin Having Increased
TITLE OF INVENTION: Bioluminescent Activity
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Jones & Askew
STREET: 191 Peachtree Street, 37th Floor
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30303-1769
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch
COMPUTER: Macintosh
OPERATING SYSTEM: 7.0
SOFTWARE: Microsoft Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/331,379
FILING DATE: October 28, 1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: No. 5541309
FILING DATE: No. 5541309
ATTORNEY/AGENT INFORMATION:
NAME: Larry W. Stults, Ph.D.
REGISTRATION NUMBER: 34,025
REFERENCE/DOCKET NUMBER: 04840-0071
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-818-3799
TELEFAX: 404-818-3799

INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 196 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Recombinant site-directed aequorin mutant
OTHER INFORMATION: Site-directed mutant having increased bioluminescent activity
PUBLICATION INFORMATION:
RELEVANT RESIDUES IN SEQ ID NO: 3: Asp 124 changed to Ser
US-08-331-379-3
Query Match 100.0%; Score 27; DB 1; Length 196;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGAVP 5
|||||
Db 192 GGAVP 196
RESULT 11
US-08-331-379-4
Sequence 4, Application US/08331379
Patent No. 5541309
GENERAL INFORMATION:
APPLICANT: Prasher, Douglas
TITLE OF INVENTION: Modified Aequorin Having Increased
TITLE OF INVENTION: Bioluminescent Activity
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Jones & Askew
STREET: 191 Peachtree Street, 37th Floor
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30303-1769
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch
COMPUTER: Macintosh
OPERATING SYSTEM: 7.0
SOFTWARE: Microsoft Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/331,379
FILING DATE: October 28, 1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: No. 5541309
FILING DATE: No. 5541309
ATTORNEY/AGENT INFORMATION:
NAME: Larry W. Stults, Ph.D.
REGISTRATION NUMBER: 34,025
REFERENCE/DOCKET NUMBER: 04840-0071
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-818-3799
TELEFAX: 404-818-3799
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 196 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Recombinant site-directed aequorin mutant
OTHER INFORMATION: Site-directed aequorin mutant having increased bioluminescent activity
PUBLICATION INFORMATION:
RELEVANT RESIDUES IN SEQ ID NO: 4: Glu 135 changed to Ser
US-08-331-379-4
Query Match 100.0%; Score 27; DB 1; Length 196;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGAVP 5
 11111
 DB 192 GGAVP 196

RESULT 12
 US-08-331-379-5

Sequence 5, Application US/08331379
 Patent No. 5541309

GENERAL INFORMATION:

APPLICANT: Prasher, Douglas

TITLE OF INVENTION: Modified Aequorin Having Increased

TITLE OF INVENTION: Bioluminescent Activity

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Jones & Askew

STREET: 191 Peachtree Street, 37th Floor

CITY: Atlanta

STATE: Georgia

COUNTRY: USA

ZIP: 30303-1769

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch

COMPUTER: Macintosh

OPERATING SYSTEM: 7.0

SOFTWARE: Microsoft Word

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/331,379

FILING DATE: October 28, 1994

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: NO. 5541309e

FILING DATE: NO. 5541309e

ATTORNEY/AGENT INFORMATION:

NAME: Larry W. Stults, Ph.D.

REGISTRATION NUMBER: 34,025

REFERENCE/DOCKET NUMBER: 04840-0071

TELECOMMUNICATION INFORMATION:

TELEPHONE: 404-818-3700

TELEFAX: 404-818-3799

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 196 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

FEATURE:

NAME/KEY: Recombinant site-directed aequorin mutant

OTHER INFORMATION: Site-directed mutant having increased bioluminescent activity

PUBLICATION INFORMATION:

RELEVANT RESIDUES IN SEQ ID NO: 5: Gly 129 changed to Ala

US-08-331-379-5

Query Match 100.0%; Score 27; DB 1; Length 196;

Best Local Similarity 100.0%; Pred. No. 2.1e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGAVP 5

DB 192 GGAVP 196

RESULT 13

US-08-583-672-2

Sequence 2, Application US/08583672

Patent No. 5741673

GENERAL INFORMATION:

APPLICANT: Montminy, Marc R.

APPLICANT: Leonard, James N.

TITLE OF INVENTION: A NOVEL HOMEOBOX FACTOR THAT STIMULATES

TITLE OF INVENTION: INSULIN EXPRESSION IN PANCREATIC ISLET CELLS

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESSES:
 ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
 STREET: 444 South Flower Street, Suite 2000
 CITY: Los Angeles
 STATE: CA
 COUNTRY: USA
 ZIP: 90071

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/583,672

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/106,936

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Reiter, Stephen E.

REGISTRATION NUMBER: 31,192

REFERENCE/DOCKET NUMBER: P41 9422

TELECOMMUNICATION INFORMATION:

TELEPHONE: 619-546-4737

TELEFAX: 619-546-9392

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 283 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-583-672-2

Query Match 100.0%; Score 27; DB 1; Length 283;

Best Local Similarity 100.0%; Pred. No. 3e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGAVP 5

DB 245 GGAVP 249

RESULT 14

US-08-320-148B-2

Sequence 2, Application US/08320148B

Patent No. 5849989

GENERAL INFORMATION:

APPLICANT: Edlund, Thomas

TITLE OF INVENTION: Insulin Promoter Factor, and Uses

TITLE OF INVENTION: Related Thereof

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESSES:

ADDRESSEE: LAHIVE & COCKFIELD

STREET: 60 State Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: ASCII (text)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/320,148B

FILING DATE: 07-OCT-1994

CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: Vincent, Matthew P.

REGISTRATION NUMBER: 36,709

REFERENCE/DOCKET NUMBER: ONI-004

TELECOMMUNICATION INFORMATION:

Search completed: March 17, 2003, 12:11:52
Job time : 13 secs

TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 284 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-320-148B-2

Query Match 100.0%; Score 27; DB 2; Length 284;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGAVP 5
|||||
DB 246 GGAVP 250

RESULT 15
US-08-589-028-6
Sequence 6, Application US/08589028
Patent No. 6087129
GENERAL INFORMATION:
APPLICANT: Newgard, Christopher B.
APPLICANT: Halban, Philippe
APPLICANT: No. 6087129minington, Karl D.
APPLICANT: Clark, Samuel A.
APPLICANT: Thijsen, Anice E.
APPLICANT: Quade, Christian
TITLE OF INVENTION: Recombinant Expression of Proteins From
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P. O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/589,028
FILING DATE: Concurrently Herewith
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Highlander, Steven L.
REGISTRATION NUMBER: 47,642
REFERENCE/DOCKET NUMBER: UTSD:426\ATL
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 284 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-589-028-6

Query Match 100.0%; Score 27; DB 3; Length 284;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGAVP 5
|||||
DB 246 GGAVP 250

GenCore version 5.1.4-P5-4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 17, 2003, 12:07:24 ; Search time 152 Seconds

(Without alignments)

21.208 Million cell updates/sec

Title: SEQ25-AV

Perfect score: 27

Sequence: 1 ggapv 5

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

pending_Patents_AA_Main:*

- 1: /cgn2_6/ptodata/2/paa/US06_COMB.pep:*
- 2: /cgn2_6/ptodata/2/paa/US06_COMB.pep:*
- 3: /cgn2_6/ptodata/2/paa/US07_COMB.pep:*
- 4: /cgn2_6/ptodata/2/paa/US08_COMB.pep:*
- 5: /cgn2_6/ptodata/2/paa/US081_COMB.pep:*
- 6: /cgn2_6/ptodata/2/paa/US082_COMB.pep:*
- 7: /cgn2_6/ptodata/2/paa/US083_COMB.pep:*
- 8: /cgn2_6/ptodata/2/paa/US084_COMB.pep:*
- 9: /cgn2_6/ptodata/2/paa/US085_COMB.pep:*
- 10: /cgn2_6/ptodata/2/paa/US086_COMB.pep:*
- 11: /cgn2_6/ptodata/2/paa/US087_COMB.pep:*
- 12: /cgn2_6/ptodata/2/paa/US088_COMB.pep:*
- 13: /cgn2_6/ptodata/2/paa/US089_COMB.pep:*
- 14: /cgn2_6/ptodata/2/paa/US090_COMB.pep:*
- 15: /cgn2_6/ptodata/2/paa/US091_COMB.pep:*
- 16: /cgn2_6/ptodata/2/paa/US092_COMB.pep:*
- 17: /cgn2_6/ptodata/2/paa/US093_COMB.pep:*
- 18: /cgn2_6/ptodata/2/paa/US094_COMB.pep:*
- 19: /cgn2_6/ptodata/2/paa/US095_COMB.pep:*
- 20: /cgn2_6/ptodata/2/paa/US096_COMB.pep:*
- 21: /cgn2_6/ptodata/2/paa/US097_COMB.pep:*
- 22: /cgn2_6/ptodata/2/paa/US098_COMB.pep:*
- 23: /cgn2_6/ptodata/2/paa/US099_COMB.pep:*
- 24: /cgn2_6/ptodata/2/paa/US100_COMB.pep:*
- 25: /cgn2_6/ptodata/2/paa/US101_COMB.pep:*
- 26: /cgn2_6/ptodata/2/paa/US102_COMB.pep:*
- 27: /cgn2_6/ptodata/2/paa/US60_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	27	100.0	10	US-09-011-545-37	Sequence 37, Appl
2	27	100.0	23	US-09-538-038-837	Sequence 837, App
3	27	100.0	44	US-60-160-189-8914	Sequence 8914, App
4	27	100.0	46	US-09-617-682A-1242	Sequence 1242, Ap
5	27	100.0	46	US-09-688-052-3054	Sequence 3054, Ap
6	27	100.0	46	US-09-855-754-8	Sequence 8, Appl

7	27	100.0	47	1	PCT-US01-01334-4311	Sequence 4311, Ap
8	27	100.0	47	21	US-09-764-874-4311	Sequence 4311, Ap
9	27	100.0	49	21	US-09-760-479-905	Sequence 905, App
10	27	100.0	49	26	US-10-206-008-905	Sequence 905, App
11	27	100.0	51	22	US-09-855-754-7	Sequence 7, Appl
12	27	100.0	51	22	US-09-855-754-13	Sequence 13, Appl
13	27	100.0	54	27	US-60-140-956-2004	Sequence 2004, Ap
14	27	100.0	56	22	US-09-855-754-9	Sequence 9, Appl
15	27	100.0	56	22	US-09-855-754-10	Sequence 10, Appl
16	27	100.0	56	22	US-09-855-754-12	Sequence 12, Appl
17	27	100.0	61	20	US-09-617-682A-10471	Sequence 10471, A
18	27	100.0	61	22	US-09-855-754-11	Sequence 11, Appl
19	27	100.0	63	27	US-60-187-385-859	Sequence 859, App
20	27	100.0	64	21	US-09-708-427-42392	Sequence 42392, A
21	27	100.0	65	1	PCT-US01-08631-58182	Sequence 58182, A
22	27	100.0	66	6	US-08-256-769B-31	Sequence 31, Appl
23	27	100.0	66	20	US-09-617-681A-7282	Sequence 7282, Ap
24	27	100.0	66	27	US-60-163-062-926	Sequence 926, App
25	27	100.0	67	21	US-09-758-450-896	Sequence 896, App
26	27	100.0	67	26	US-10-227-582-896	Sequence 896, App
27	27	100.0	69	20	US-09-620-111B-6519	Sequence 6519, App
28	27	100.0	73	20	US-09-688-052-3212	Sequence 3212, Ap
29	27	100.0	76	19	US-09-513-996A-79827	Sequence 79827, A
30	27	100.0	79	27	US-60-160-203-5975	Sequence 5975, Ap
31	27	100.0	79	27	US-60-169-840-9002	Sequence 9002, Ap
32	27	100.0	79	27	US-60-187-385-901	Sequence 901, App
33	27	100.0	80	20	US-09-617-681A-192	Sequence 192, App
34	27	100.0	81	21	US-09-708-427-75445	Sequence 75445, A
35	27	100.0	82	20	US-09-617-681A-6611	Sequence 6611, Ap
36	27	100.0	83	20	US-09-617-681A-4459	Sequence 4459, Ap
37	27	100.0	84	20	US-09-617-682A-1241	Sequence 1241, Ap
38	27	100.0	85	27	US-60-207-359-238	Sequence 238, App
39	27	100.0	86	24	US-10-029-386-30668	Sequence 30668, A
40	27	100.0	87	20	US-09-688-052-3058	Sequence 3058, Ap
41	27	100.0	90	20	US-09-620-111B-6517	Sequence 6517, Ap
42	27	100.0	91	21	US-09-617-681A-7539	Sequence 7539, Ap
43	27	100.0	91	21	US-09-791-537-34282	Sequence 34282, A
44	27	100.0	91	21	US-09-791-537-63331	Sequence 63331, A
45	27	100.0	92	26	US-10-219-999-39144	Sequence 39144, A

ALIGNMENTS

RESULT 1
US-09-011-545-37
Sequence 37, Application US/09011545C
GENERAL INFORMATION:
APPLICANT: SUMITOMO ELECTRIC INDUSTRIES, LTD.
TITLE OF INVENTION: BACTERIAL PEPTIDE LIBRARY
FILE REFERENCE: 7898/245156
CURRENT APPLICATION NUMBER: US/09/011,545C
CURRENT FILING DATE: 1998-05-20
PRIOR FILING DATE: JP 7/199745
NUMBER OF SEQ ID NOS: 78
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 37
LENGTH: 10
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: The peptide encoded by the DNA set forth in SEQ ID NO: 15
US-09-011-545-37

Query Match 100.0%; Score 27; DB 14; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 GGAVP 5
DB 2 GGAVP 6

```

RESULT 2
US-09-538-038-837
; Sequence 837, Application US/09538038
; GENERAL INFORMATION:
; APPLICANT: Beasley, James
; APPLICANT: Blume, Arthur J.
; APPLICANT: Schaffer, Laugel
; APPLICANT: Pillutla, Renuka
; APPLICANT: Brandt, Jakob
; APPLICANT: Brissette, Renee
; APPLICANT: Spetzler, Jane
; APPLICANT: Cheng, Weiqing
; APPLICANT: Ostergaard, Soren
; APPLICANT: Mandeckl, Wlodek S.
; APPLICANT: Hansen, Per Hertz
; APPLICANT: Ravera, Mark
; APPLICANT: Hsiao, Ku-chuan
; TITLE OF INVENTION: INSULIN AND IGF-1 RECEPTOR AGONISTS AND ANTAGONISTS
; FILE REFERENCE: 1878-4051
; CURRENT APPLICATION NUMBER: US/09/538,038
; CURRENT FILING DATE: 2000-03-29
; NUMBER OF SEQ ID NOS: 2653
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 837
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: peptide
US-09-538-038-837

```

```

Query Match
Best Local Similarity 100.0%; Score 27; DB 19; Length 23;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGAVP 5
DB 2 GGAVP 6

```

```

RESULT 3
US-60-160-189-8914
; Sequence 8914, Application US/60160189
; GENERAL INFORMATION:
; APPLICANT: BONAZZI, VIVIAN
; TITLE OF INVENTION: ISOLATED HUMAN DRUG TARGET PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN DRUG TARGET PROTEINS
; FILE REFERENCE: C1000112
; CURRENT APPLICATION NUMBER: US/60/160,189
; CURRENT FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 10162
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8914
; LENGTH: 44
; TYPE: PRT
; ORGANISM: HUMAN
US-60-160-189-8914

```

```

Query Match
Best Local Similarity 100.0%; Score 27; DB 27; Length 44;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGAVP 5
DB 6 GGAVP 10

```

```

RESULT 4
US-09-617-682A-1242
; Sequence 1242, Application US/09617682A

```

```

; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai
; APPLICANT: BROVER, Vyacheslav
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptide
; TITLE OF INVENTION: Thereby
; FILE REFERENCE: 2750-1063P
; CURRENT APPLICATION NUMBER: US/09/617,682A
; CURRENT FILING DATE: 2000-07-19
; NUMBER OF SEQ ID NOS: 16871
; SEQ ID NO 1242
; LENGTH: 46
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..46
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..46
; OTHER INFORMATION: Ceres Seq. ID 1345818
US-09-617-682A-1242

```

```

Query Match
Best Local Similarity 100.0%; Score 27; DB 20; Length 46;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGAVP 5
DB 29 GGAVP 33

```

```

RESULT 5
US-09-688-052-3054
; Sequence 3054, Application US/09688052
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai
; APPLICANT: BROVER, Vyacheslav
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptide
; FILE REFERENCE: 2750-1238P
; CURRENT APPLICATION NUMBER: US/09/688,052
; CURRENT FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 4086
; SEQ ID NO 3054
; LENGTH: 46
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..46
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..46
; OTHER INFORMATION: Ceres Seq. ID 1599998
US-09-688-052-3054

```

```

Query Match
Best Local Similarity 100.0%; Score 27; DB 20; Length 46;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGAVP 5
DB 29 GGAVP 33

```

```

RESULT 6
US-09-855-754-8
; Sequence 8, Application US/09855754
; GENERAL INFORMATION:
; APPLICANT: BOURSIAUX-EUDE, CAROLINE
; APPLICANT: GUISSO-MACLOUF, NICOLE
; TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED
; TITLE OF INVENTION: REGIONS OF PERTACTIN IN BORDETELLA PERTUSSIS,

```

```

: TITLE OF INVENTION: BORDETELLA PARAPERTUSSIS AND BORDETELLA
: TITLE OF INVENTION: BRONCHISEPTICA, THEIR USE IN DIAGNOSTICS, AND IN
: FILE REFERENCE: 03495-0206-00000
: CURRENT FILING DATE: 2001-09-10
: PRIOR APPLICATION NUMBER: 60/206,969
: PRIOR FILING DATE: 2000-05-25
: SOFTWARE: Patent Ver. 2.1
: SEQ ID NO 8
: LENGTH: 46
: TYPE: PRT
: ORGANISM: Bordetella bronchiseptica
: US-09-855-754-8

Query Match      100.0%; Score 27; DB 22; Length 46;
Best Local Similarity 100.0%; Pred. No. 8.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GGAVP 5
Db      13 GGAVP 17

RESULT 7
PCT-US01-01334-4311
: Sequence 4311, Application PC/TUS0101334
: GENERAL INFORMATION:
: APPLICANT: Human Genome Sciences, Inc., et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
: FILE REFERENCE: PC010PCT
: CURRENT APPLICATION NUMBER: PCT/US01/01334
: PRIOR FILING DATE: 2001-01-17
: PRIOR APPLICATION NUMBER: 60/179,065
: PRIOR FILING DATE: 2000-01-31
: PRIOR APPLICATION NUMBER: 60/180,628
: PRIOR FILING DATE: 2000-02-04
: PRIOR APPLICATION NUMBER: 60/214,886
: PRIOR FILING DATE: 2000-06-28
: PRIOR APPLICATION NUMBER: 60/217,487
: PRIOR FILING DATE: 2000-07-11
: PRIOR APPLICATION NUMBER: 60/225,758
: PRIOR FILING DATE: 2000-08-14
: PRIOR APPLICATION NUMBER: 60/220,963
: PRIOR FILING DATE: 2000-07-26
: PRIOR APPLICATION NUMBER: 60/217,496
: PRIOR FILING DATE: 2000-07-11
: PRIOR APPLICATION NUMBER: 60/225,447
: PRIOR FILING DATE: 2000-08-14
: PRIOR APPLICATION NUMBER: 60/218,290
: PRIOR FILING DATE: 2000-07-14
: PRIOR APPLICATION NUMBER: 60/225,757
: PRIOR FILING DATE: 2000-08-14
: PRIOR APPLICATION NUMBER: 60/226,868
: PRIOR FILING DATE: 2000-08-22
: PRIOR APPLICATION NUMBER: 60/216,647
: PRIOR FILING DATE: 2000-07-07
: PRIOR APPLICATION NUMBER: 60/225,267
: PRIOR FILING DATE: 2000-08-14
: PRIOR APPLICATION NUMBER: 60/216,880
: PRIOR FILING DATE: 2000-07-07
: PRIOR APPLICATION NUMBER: 60/225,270
: PRIOR FILING DATE: 2000-08-14
: PRIOR APPLICATION NUMBER: 60/251,869
: PRIOR FILING DATE: 2000-12-08
: PRIOR APPLICATION NUMBER: 60/225,834
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: 60/234,274
: PRIOR FILING DATE: 2000-09-21
: PRIOR APPLICATION NUMBER: 60/234,223
: PRIOR FILING DATE: 2000-09-21
: PRIOR APPLICATION NUMBER: 60/228,924

: PRIOR FILING DATE: 2000-08-30
: PRIOR APPLICATION NUMBER: 60/224,518
: PRIOR FILING DATE: 2000-08-14
: PRIOR APPLICATION NUMBER: 60/236,369
: PRIOR FILING DATE: 2000-09-29
: PRIOR APPLICATION NUMBER: 60/224,519
: PRIOR FILING DATE: 2000-08-14
: PRIOR APPLICATION NUMBER: 60/220,964
: PRIOR FILING DATE: 2000-07-26
: PRIOR APPLICATION NUMBER: 60/241,809
: PRIOR FILING DATE: 2000-10-20
: PRIOR APPLICATION NUMBER: 60/249,299
: PRIOR FILING DATE: 2000-11-17
: PRIOR APPLICATION NUMBER: 60/236,327
: PRIOR FILING DATE: 2000-09-29
: PRIOR APPLICATION NUMBER: 60/241,785
: PRIOR FILING DATE: 2000-10-20
: PRIOR APPLICATION NUMBER: 60/244,617
: PRIOR FILING DATE: 2000-11-01
: PRIOR APPLICATION NUMBER: 60/225,268
: PRIOR FILING DATE: 2000-08-14
: PRIOR APPLICATION NUMBER: 60/236,368
: PRIOR FILING DATE: 2000-09-29
: PRIOR APPLICATION NUMBER: 60/251,856
: PRIOR FILING DATE: 2000-12-08
: PRIOR APPLICATION NUMBER: 60/251,868
: PRIOR FILING DATE: 2000-12-08
: PRIOR APPLICATION NUMBER: 60/229,344
: PRIOR FILING DATE: 2000-09-01
: PRIOR APPLICATION NUMBER: 60/234,997
: PRIOR FILING DATE: 2000-09-25
: PRIOR APPLICATION NUMBER: 60/229,343
: PRIOR FILING DATE: 2000-09-01
: PRIOR APPLICATION NUMBER: 60/229,345
: PRIOR FILING DATE: 2000-09-01
: PRIOR APPLICATION NUMBER: 60/229,287
: PRIOR FILING DATE: 2000-09-01
: PRIOR APPLICATION NUMBER: 60/229,513
: PRIOR FILING DATE: 2000-09-05
: PRIOR APPLICATION NUMBER: 60/223,413
: PRIOR FILING DATE: 2000-09-08
: PRIOR APPLICATION NUMBER: 60/229,509
: PRIOR FILING DATE: 2000-09-05
: PRIOR APPLICATION NUMBER: 60/236,367
: PRIOR FILING DATE: 2000-09-29
: PRIOR APPLICATION NUMBER: 60/237,039
: PRIOR FILING DATE: 2000-10-02
: PRIOR APPLICATION NUMBER: 60/237,038
: PRIOR FILING DATE: 2000-10-02
: PRIOR APPLICATION NUMBER: 60/236,370
: PRIOR FILING DATE: 2000-09-29
: PRIOR APPLICATION NUMBER: 60/236,802
: PRIOR FILING DATE: 2000-10-02
: PRIOR APPLICATION NUMBER: 60/237,037
: PRIOR FILING DATE: 2000-10-02
: PRIOR APPLICATION NUMBER: 60/237,040
: PRIOR FILING DATE: 2000-10-02
: PRIOR APPLICATION NUMBER: 60/240,960
: PRIOR FILING DATE: 2000-10-20
: PRIOR APPLICATION NUMBER: 60/239,935
: PRIOR FILING DATE: 2000-10-13
: PRIOR APPLICATION NUMBER: 60/239,937
: PRIOR FILING DATE: 2000-10-13
: PRIOR APPLICATION NUMBER: 60/241,787
: PRIOR FILING DATE: 2000-10-20
: PRIOR APPLICATION NUMBER: 60/246,474
: PRIOR FILING DATE: 2000-11-08
: PRIOR APPLICATION NUMBER: 60/246,532
: PRIOR FILING DATE: 2000-11-08
: PRIOR APPLICATION NUMBER: 60/249,216
: PRIOR FILING DATE: 2000-11-17
: PRIOR APPLICATION NUMBER: 60/249,210
: PRIOR FILING DATE: 2000-11-17
```

```

PRIOR APPLICATION NUMBER: 60/226,681
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 60/225,759
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/225,213
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/227,182
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 60/225,214
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/235,836
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/230,438
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/215,135
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: 60/225,266
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/249,218
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,208
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,213
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,212
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,207
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,245
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,244
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,217
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,211
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,215
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,264
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,214
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,297
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/232,400
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/231,242
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/232,081
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/232,080
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/231,414
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/231,244
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/233,064
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/233,063
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/232,397
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/232,399
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/232,401
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/241,808
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/241,826
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/241,786
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/241,221

```

```

PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/246,475
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 60/231,243
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/233,065
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/232,398

Query Match
Best Local Similarity 100.0%; Score 27; DB 1; Length 47;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAVP 5
DB 42 GGAVP 46

RESULT 8
US-09-764-874-4311
Sequence 4311, Application US/09764874
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PC010
CURRENT FILING DATE: 2001-01-17
PRIOR APPLICATION NUMBER: 60/179,065
PRIOR FILING DATE: 2000-01-31
PRIOR APPLICATION NUMBER: 60/180,628
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: 60/214,886
PRIOR FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: 60/217,487
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/225,758
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/220,963
PRIOR FILING DATE: 2000-07-26
PRIOR APPLICATION NUMBER: 60/217,496
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/225,447
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/218,290
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: 60/225,757
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/226,868
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 60/216,647
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: 60/225,267
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/216,880
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: 60/225,270
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/251,869
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/235,834
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/234,274
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: 60/234,223
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: 60/228,924
PRIOR FILING DATE: 2000-08-30
PRIOR APPLICATION NUMBER: 60/224,518
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/236,369
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/224,519
PRIOR FILING DATE: 2000-08-14

```

PRIOR APPLICATION NUMBER: 60/220,964
PRIOR FILING DATE: 2000-07-26
PRIOR APPLICATION NUMBER: 60/241,809
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/249,299
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/236,327
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/241,785
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/244,617
PRIOR FILING DATE: 2000-11-01
PRIOR APPLICATION NUMBER: 60/225,268
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/236,368
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/251,856
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/251,868
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/229,344
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/234,997
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: 60/229,343
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229,345
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229,287
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229,513
PRIOR FILING DATE: 2000-09-05
PRIOR APPLICATION NUMBER: 60/231,413
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/229,509
PRIOR FILING DATE: 2000-09-05
PRIOR APPLICATION NUMBER: 60/236,367
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/237,039
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237,038
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/236,370
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/236,802
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237,037
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237,040
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/240,960
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/239,935
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 60/239,937
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 60/241,787
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/246,474
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 60/246,532
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 60/249,216
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,210
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/226,681
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 60/225,759
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/225,213
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/227,182

PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 60/225,214
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/235,836
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/230,438
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/215,135
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: 60/225,266
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/249,218
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,208
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,213
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,212
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,207
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,245
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,244
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,217
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,211
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,215
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,264
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,214
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,297
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/232,400
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/231,242
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/232,081
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/232,080
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/231,414
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/231,244
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/233,064
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/233,063
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/232,397
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/232,399
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/232,401
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/241,808
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/241,826
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/241,786
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/241,221
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/246,475
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 60/231,243
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/233,065
PRIOR FILING DATE: 2000-09-14

```

; PRIOR APPLICATION NUMBER: 60/232,398
Query Match      100.0%; Score 27; DB 21; Length 47;
Best Local Similarity 100.0%; Pred. No. 8.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GGAVP 5
Db      42 GGAVP 46

RESULT 9
US-09-760-479-905
; Sequence 905, Application US/09760479
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT253
; CURRENT APPLICATION NUMBER: US/09/760,479
; CURRENT FILING DATE: 2001-01-16
; PRIOR APPLICATION data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 946
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 905
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (29)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-760-479-905

Query Match      100.0%; Score 27; DB 21; Length 49;
Best Local Similarity 100.0%; Pred. No. 9.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GGAVP 5
Db      32 GGAVP 36

RESULT 10
US-10-206-008-905
; Sequence 905, Application US/10206008
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT253CIN
; CURRENT APPLICATION NUMBER: US/10/206,008
; CURRENT FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 09/760,479
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/225,757
; PRIOR FILING DATE: 2000-08-14

```

```

PRIOR APPLICATION NUMBER: 60/226,868
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 60/216,647
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: 60/225,267
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/216,880
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: 60/225,270
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/251,869
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/235,834
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/234,274
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: 60/234,223
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: 60/228,924
PRIOR FILING DATE: 2000-08-30
PRIOR APPLICATION NUMBER: 60/224,518
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/236,369
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/224,519
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/220,964
PRIOR FILING DATE: 2000-07-26
PRIOR APPLICATION NUMBER: 60/241,809
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/249,299
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/236,327
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/241,785
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/244,617
PRIOR FILING DATE: 2000-11-01
PRIOR APPLICATION NUMBER: 60/225,268
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/236,368
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/251,856
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/251,868
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/229,344
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/234,997
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: 60/229,343
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229,345
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229,287
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229,513
PRIOR FILING DATE: 2000-09-05
PRIOR APPLICATION NUMBER: 60/231,413
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/229,509
PRIOR FILING DATE: 2000-09-05
PRIOR APPLICATION NUMBER: 60/236,367
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/237,039
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237,038
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/236,370
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/236,802
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237,037

```


;; PRIOR FILING DATE: 2000-10-02
;; PRIOR APPLICATION NUMBER: 60/237,040
;; PRIOR FILING DATE: 2000-10-02
;; PRIOR APPLICATION NUMBER: 60/240,960
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/239,935
;; PRIOR FILING DATE: 2000-10-13
;; PRIOR APPLICATION NUMBER: 60/239,937
;; PRIOR FILING DATE: 2000-10-13
;; PRIOR APPLICATION NUMBER: 60/241,787
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/246,474
;; PRIOR FILING DATE: 2000-11-08
;; PRIOR APPLICATION NUMBER: 60/246,532
;; PRIOR FILING DATE: 2000-11-08
;; PRIOR APPLICATION NUMBER: 60/249,216
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,210
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/226,681
;; PRIOR FILING DATE: 2000-08-22
;; PRIOR APPLICATION NUMBER: 60/225,759
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/225,213
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/227,182
;; PRIOR FILING DATE: 2000-08-22
;; PRIOR APPLICATION NUMBER: 60/225,214
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/235,836
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: 60/230,438
;; PRIOR FILING DATE: 2000-09-06
;; PRIOR APPLICATION NUMBER: 60/215,135
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: 60/225,266
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/249,218
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,208
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,213
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,212
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,207
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,245
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,244
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,217
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,211
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,215
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,264
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,214
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,297
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/232,400
;; PRIOR FILING DATE: 2000-09-14
;; PRIOR APPLICATION NUMBER: 60/231,242
;; PRIOR FILING DATE: 2000-09-08
;; PRIOR APPLICATION NUMBER: 60/232,081
;; PRIOR FILING DATE: 2000-09-08
;; PRIOR APPLICATION NUMBER: 60/232,080
;; PRIOR FILING DATE: 2000-09-08
;; PRIOR APPLICATION NUMBER: 60/231,414
;; PRIOR FILING DATE: 2000-09-08

;; PRIOR APPLICATION NUMBER: 60/231,244
;; PRIOR FILING DATE: 2000-09-08
;; PRIOR APPLICATION NUMBER: 60/233,064
;; PRIOR FILING DATE: 2000-09-14
;; PRIOR APPLICATION NUMBER: 60/233,063
;; PRIOR FILING DATE: 2000-09-14
;; PRIOR APPLICATION NUMBER: 60/232,397
;; PRIOR FILING DATE: 2000-09-14
;; PRIOR APPLICATION NUMBER: 60/232,399
;; PRIOR FILING DATE: 2000-09-14
;; PRIOR APPLICATION NUMBER: 60/232,401
;; PRIOR FILING DATE: 2000-09-14
;; PRIOR APPLICATION NUMBER: 60/241,808
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/241,826
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/241,786
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/241,221
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/246,475
;; PRIOR FILING DATE: 2000-11-08
;; PRIOR APPLICATION NUMBER: 60/231,243
;; PRIOR FILING DATE: 2000-09-08
;; PRIOR APPLICATION NUMBER: 60/233,065

Query Match 100.0%; Score 27; DB 26; Length 49;
Best Local Similarity 100.0%; Pred. No. 9, 6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GGAVP 5
Db 32 GGAVP 36

RESULT 11
US-09-855-754-7
;; Sequence 7, Application US/09855754
;; GENERAL INFORMATION:
;; APPLICANT: BOURSAX- EUDE, CAROLINE
;; TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED
;; TITLE OF INVENTION: REGIONS OF PERTACTIN IN BORDETELLA PERTUSSIS,
;; TITLE OF INVENTION: BORDETELLA PARAPERTUSSIS, AND BORDETELLA
;; TITLE OF INVENTION: BRONCHISEPTICA, THEIR USE IN DIAGNOSTICS, AND IN
;; FILE REFERENCE: 03495-0206-00000
;; CURRENT APPLICATION NUMBER: US/09/855,754
;; CURRENT FILING DATE: 2001-09-10
;; PRIOR APPLICATION NUMBER: 60/206,969
;; PRIOR FILING DATE: 2000-05-25
;; NUMBER OF SEQ ID NOS: 24
;; SOFTWARE: Patentln Ver. 2.1
;; SEQ ID NO 7
;; LENGTH: 51
;; TYPE: PRT
;; ORGANISM: Bordetella bronchiseptica
US-09-855-754-7

Query Match 100.0%; Score 27; DB 22; Length 51;
Best Local Similarity 100.0%; Pred. No. 9, 6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GGAVP 5
Db 13 GGAVP 17

RESULT 12
US-09-855-754-13
;; Sequence 13, Application US/09855754
;; GENERAL INFORMATION:
;; APPLICANT: BOURSAX- EUDE, CAROLINE

```

; APPLICANT: GUIISO-MACLOUF, NICOLE
; TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED
; TITLE OF INVENTION: REGIONS OF PERTACTIN IN BORDETELLA PERTUSSIS,
; TITLE OF INVENTION: BORDETELLA PARAPERTUSSIS, AND BORDETELLA
; TITLE OF INVENTION: BRONCHISEPTICA, THEIR USE IN DIAGNOSTICS, AND IN
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS
; FILE REFERENCE: 03495-0206-00000
; CURRENT APPLICATION NUMBER: US/09/855,754
; CURRENT FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/206,969
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 51
; TYPE: PRT
; ORGANISM: Bordetella bronchiseptica
US-09-855-754-13

```

```

Query Match
Best Local Similarity 100.0%; Score 27; DB 22; Length 51;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 GGAVP 5
    |||||
DB 13 GGAVP 17

```

```

RESULT 13
US-60-140-956-2004
; Sequence 2004, Application US/60140956
; GENERAL INFORMATION:
; APPLICANT: Keriavage, Anthony
; TITLE OF INVENTION: ISOLATED SECRETED PROTEINS, NUCLEIC ACID
; TITLE OF INVENTION: MOLECULES ENCODING SECRETED PROTEINS, AND USES THEREOF
; FILE REFERENCE: CLO00041
; CURRENT APPLICATION NUMBER: US/60/140,956
; CURRENT FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 2638
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2004
; LENGTH: 54
; TYPE: PRT
; ORGANISM: Drosophila
US-60-140-956-2004

```

```

Query Match
Best Local Similarity 100.0%; Score 27; DB 27; Length 54;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 GGAVP 5
    |||||
DB 40 GGAVP 44

```

```

RESULT 14
US-09-855-754-9
; Sequence 9, Application US/09855754
; GENERAL INFORMATION:
; APPLICANT: BOURSAX-EUDE, CAROLINE
; TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED
; TITLE OF INVENTION: REGIONS OF PERTACTIN IN BORDETELLA PERTUSSIS,
; TITLE OF INVENTION: BORDETELLA PARAPERTUSSIS, AND BORDETELLA
; TITLE OF INVENTION: BRONCHISEPTICA, THEIR USE IN DIAGNOSTICS, AND IN
; FILE REFERENCE: 03495-0206-00000
; CURRENT APPLICATION NUMBER: US/09/855,754
; CURRENT FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/206,969
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1

```

```

; SEQ ID NO 9
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Bordetella bronchiseptica
US-09-855-754-9

```

```

Query Match
Best Local Similarity 100.0%; Score 27; DB 22; Length 56;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 GGAVP 5
    |||||
DB 18 GGAVP 22

```

```

RESULT 15
US-09-855-754-10
; Sequence 10, Application US/09855754
; GENERAL INFORMATION:
; APPLICANT: BOURSAX-EUDE, CAROLINE
; APPLICANT: GUIISO-MACLOUF, NICOLE
; TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED
; TITLE OF INVENTION: REGIONS OF PERTACTIN IN BORDETELLA PERTUSSIS,
; TITLE OF INVENTION: BORDETELLA PARAPERTUSSIS, AND BORDETELLA
; TITLE OF INVENTION: BRONCHISEPTICA, THEIR USE IN DIAGNOSTICS, AND IN
; FILE REFERENCE: 03495-0206-00000
; CURRENT APPLICATION NUMBER: US/09/855,754
; CURRENT FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/206,969
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Bordetella bronchiseptica
US-09-855-754-10

```

```

Query Match
Best Local Similarity 100.0%; Score 27; DB 22; Length 56;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 GGAVP 5
    |||||
DB 13 GGAVP 17

```

```

Search completed: March 17, 2003, 12:19:38
Job time : 154 secs

```

GenCore version 5.1.4.P5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 17, 2003, 12:09:04 ; Search time 23 Seconds
(without alignments)
28.366 Million cell updates/sec

Title: SEQ25-AV
Perfect score: 27
Sequence: 1 ggapv 5

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 603988 seqs, 130485580 residues

Total number of hits satisfying chosen parameters: 603988

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/1/paa/PCN_NEM_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_NEM_COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_NEM_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US08_NEM_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US09_NEM_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US10_NEM_COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US80_NEM_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query length	DB ID	Description
1	27	100.0	46 5 US-09-855-754B-8	Sequence 8, Appl1
2	27	100.0	51 5 US-09-855-754B-7	Sequence 7, Appl1
3	27	100.0	51 5 US-09-855-754B-13	Sequence 13, Appl1
4	27	100.0	56 5 US-09-855-754B-9	Sequence 9, Appl1
5	27	100.0	56 5 US-09-855-754B-10	Sequence 10, Appl1
6	27	100.0	56 5 US-09-855-754B-12	Sequence 12, Appl1
7	27	100.0	59 1 PCT-US02-32727-16495	Sequence 16495, A
8	27	100.0	59 5 US-09-978-825-16495	Sequence 16495, A
9	27	100.0	59 6 US-10-057-498-16495	Sequence 16495, A
10	27	100.0	61 5 US-09-855-754B-11	Sequence 11, Appl1
11	27	100.0	84 1 PCT-US02-32727-19550	Sequence 19550, A
12	27	100.0	84 6 US-10-057-498-19550	Sequence 19550, A
13	27	100.0	84 6 US-10-057-498-19550	Sequence 19550, A
14	27	100.0	105 1 PCT-US02-32727-27641	Sequence 27641, A
15	27	100.0	105 5 US-09-978-825-27641	Sequence 27641, A
16	27	100.0	105 6 US-10-057-498-27641	Sequence 27641, A
17	27	100.0	109 6 US-10-366-683-18907	Sequence 18907, A
18	27	100.0	119 6 US-10-366-683-22545	Sequence 22545, A
19	27	100.0	142 6 US-10-218-140-834	Sequence 834, App
20	27	100.0	148 1 PCT-US02-32727-7685	Sequence 7685, Ap
21	27	100.0	148 5 US-09-978-825-7685	Sequence 7685, Ap
22	27	100.0	148 6 US-10-057-498-7685	Sequence 7685, Ap
23	27	100.0	149 6 US-10-366-683-22324	Sequence 22324, A
24	27	100.0	157 6 US-10-156-761-10615	Sequence 10615, A
25	27	100.0	176 1 PCT-US02-32727-26035	Sequence 26035, A
26	27	100.0	176 5 US-09-978-825-26035	Sequence 26035, A

27	27	100.0	176 6 US-10-057-498-26035	Sequence 26035, A
28	27	100.0	185 6 US-10-366-683-30603	Sequence 30603, A
29	27	100.0	187 1 PCT-US02-32727-10551	Sequence 10551, A
30	27	100.0	187 5 US-09-978-825-10551	Sequence 10551, A
31	27	100.0	187 6 US-10-057-498-10551	Sequence 10551, A
32	27	100.0	196 6 US-10-280-911-2	Sequence 2, Appl1
33	27	100.0	196 6 PCT-US02-32727-3653	Sequence 2, Appl1
34	27	100.0	199 1 PCT-US02-32727-3653	Sequence 3653, Ap
35	27	100.0	199 5 US-09-978-825-3653	Sequence 3653, Ap
36	27	100.0	199 6 US-10-057-498-3653	Sequence 3653, Ap
37	27	100.0	204 6 US-10-366-683-32993	Sequence 32993, A
38	27	100.0	211 6 US-10-366-683-32460	Sequence 32460, A
39	27	100.0	216 6 US-10-366-683-32597	Sequence 32597, A
40	27	100.0	220 6 US-10-366-683-17151	Sequence 17151, A
41	27	100.0	222 6 US-10-156-761-9484	Sequence 9484, Ap
42	27	100.0	225 7 US-60-423-586-130	Sequence 130, App
43	27	100.0	232 7 US-60-427-194-130	Sequence 130, App
44	27	100.0	235 6 US-10-366-683-23718	Sequence 23718, A
45	27	100.0	236 6 US-10-366-683-32888	Sequence 32888, A

ALIGNMENTS

RESULT 1

US-09-855-754B-8

Sequence 8, Application US/09855754B

GENERAL INFORMATION:

APPLICANT: BOURSAX-ETUDE, CAROLINE

TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED

TITLE OF INVENTION: REGIONS OF PERTACTIN IN BORDETELLA PERTUSSIS,

TITLE OF INVENTION: BORDETELLA PARAPERTUSSIS, AND BORDETELLA

TITLE OF INVENTION: BRONCHISEPTICA, THEIR USE IN DIAGNOSTICS, AND IN

FILE REFERENCE: 03495-0206-00000

CURRENT FILING DATE: 2001-05-16

PRIOR APPLICATION NUMBER: 60/206,969

NUMBER OF SEQ ID NOS: 25

SOFTWARE: Patentln Ver. 2.1

SEQ ID NO 8

LENGTH: 46

TYPE: PRT

ORGANISM: Bordetella bronchiseptica

US-09-855-754B-8

Query Match

Best Local Similarity 100.0%; Score 27; DB 5; Length 46;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAAP 5

DB 13 GGAAP 17

RESULT 2

US-09-855-754B-7

Sequence 7, Application US/09855754B

GENERAL INFORMATION:

APPLICANT: BOURSAX-ETUDE, NICOLE

TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED

TITLE OF INVENTION: REGIONS OF PERTACTIN IN BORDETELLA PERTUSSIS,

TITLE OF INVENTION: BORDETELLA PARAPERTUSSIS, AND BORDETELLA

FILE REFERENCE: 03495-0206-00000

CURRENT FILING DATE: 2001-05-16

PRIOR APPLICATION NUMBER: 60/206,969

PRIOR FILING DATE: 2000-05-25

NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 7
LENGTH: 51
TYPE: PRT
ORGANISM: Bordetella bronchiseptica
US-09-855-754B-7

Query Match 100.0%; Score 27; DB 5; Length 51;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGAVP 5
|||||
DB 13 GGAVP 17

RESULT 3
US-09-855-754B-13
Sequence 13, Application US/09855754B
GENERAL INFORMATION:
APPLICANT: BOURSAX-ENDE, CAROLINE
TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED
TITLE OF INVENTION: REGIONS OF PERTACTIN IN BORDETELLA PERTUSSIS,
TITLE OF INVENTION: BORDETELLA PARAPERTUSSIS, AND BORDETELLA
TITLE OF INVENTION: BRONCHISEPTICA, THEIR USE IN DIAGNOSTICS, AND IN
FILE REFERENCE: 03495-0206-00000
CURRENT APPLICATION NUMBER: US/09/855,754B
CURRENT FILING DATE: 2001-05-16
PRIOR APPLICATION NUMBER: 60/206,969
PRIOR FILING DATE: 2000-05-25
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 13
LENGTH: 51
TYPE: PRT
ORGANISM: Bordetella bronchiseptica
US-09-855-754B-13

Query Match 100.0%; Score 27; DB 5; Length 51;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGAVP 5
|||||
DB 13 GGAVP 17

RESULT 4
US-09-855-754B-9
Sequence 9, Application US/09855754B
GENERAL INFORMATION:
APPLICANT: BOURSAX-ENDE, CAROLINE
TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED
TITLE OF INVENTION: REGIONS OF PERTACTIN IN BORDETELLA PERTUSSIS,
TITLE OF INVENTION: BORDETELLA PARAPERTUSSIS, AND BORDETELLA
TITLE OF INVENTION: BRONCHISEPTICA, THEIR USE IN DIAGNOSTICS, AND IN
FILE REFERENCE: 03495-0206-00000
CURRENT APPLICATION NUMBER: US/09/855,754B
CURRENT FILING DATE: 2001-05-16
PRIOR APPLICATION NUMBER: 60/206,969
PRIOR FILING DATE: 2000-05-25
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 9
LENGTH: 56
TYPE: PRT
ORGANISM: Bordetella bronchiseptica
US-09-855-754B-9

Query Match 100.0%; Score 27; DB 5; Length 56;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGAVP 5
|||||
DB 18 GGAVP 22

RESULT 5
US-09-855-754B-10
Sequence 10, Application US/09855754B
GENERAL INFORMATION:
APPLICANT: BOURSAX-ENDE, CAROLINE
TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED
TITLE OF INVENTION: REGIONS OF PERTACTIN IN BORDETELLA PERTUSSIS,
TITLE OF INVENTION: BORDETELLA PARAPERTUSSIS, AND BORDETELLA
TITLE OF INVENTION: BRONCHISEPTICA, THEIR USE IN DIAGNOSTICS, AND IN
FILE REFERENCE: 03495-0206-00000
CURRENT APPLICATION NUMBER: US/09/855,754B
CURRENT FILING DATE: 2001-05-16
PRIOR APPLICATION NUMBER: 60/206,969
PRIOR FILING DATE: 2000-05-25
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 10
LENGTH: 56
TYPE: PRT
ORGANISM: Bordetella bronchiseptica
US-09-855-754B-10

Query Match 100.0%; Score 27; DB 5; Length 56;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGAVP 5
|||||
DB 13 GGAVP 17

RESULT 6
US-09-855-754B-12
Sequence 12, Application US/09855754B
GENERAL INFORMATION:
APPLICANT: BOURSAX-ENDE, CAROLINE
TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED
TITLE OF INVENTION: REGIONS OF PERTACTIN IN BORDETELLA PERTUSSIS,
TITLE OF INVENTION: BORDETELLA PARAPERTUSSIS, AND BORDETELLA
TITLE OF INVENTION: BRONCHISEPTICA, THEIR USE IN DIAGNOSTICS, AND IN
FILE REFERENCE: 03495-0206-00000
CURRENT APPLICATION NUMBER: US/09/855,754B
CURRENT FILING DATE: 2001-05-16
PRIOR APPLICATION NUMBER: 60/206,969
PRIOR FILING DATE: 2000-05-25
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 12
LENGTH: 56
TYPE: PRT
ORGANISM: Bordetella bronchiseptica
US-09-855-754B-12

Query Match 100.0%; Score 27; DB 5; Length 56;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGAVP 5
|||||

Db 13 GGAVP 17

PCT-US02-32727-16495

; Sequence 16495, Application PC/TUS0232727

; GENERAL INFORMATION:

; APPLICANT: Mitcham, Jennifer

; APPLICANT: Skeiky, Yasir

; APPLICANT: Persing, David

; APPLICANT: Bhatia, Ajay

; APPLICANT: Maisonneuve, Jean Francois

; APPLICANT: Zhang, Yanni

; APPLICANT: Wang, Siqing

; APPLICANT: Jen, Shiyun

; APPLICANT: Lodes, Michael

; APPLICANT: Benson, Darin

; APPLICANT: Jones, Robert

; APPLICANT: Carter, Darlick

; APPLICANT: Barth, Brenda

; APPLICANT: Douglass, John

; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne

; FILE REFERENCE: 210121.514C1

; CURRENT APPLICATION NUMBER: PCT/US02/32727

; CURRENT FILING DATE: 2002-10-11

; NUMBER OF SEQ ID NOS: 30992

; SEQ ID NO 16495

; LENGTH: 59

; TYPE: PRT

; ORGANISM: Propionl acnes

PCT-US02-32727-16495

Query Match 100.0%; Score 27; DB 1; Length 59;

Best Local Similarity 100.0%; Pred. No. 1.2e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGAVP 5

Db 53 GGAVP 57

RESULT 8

US-09-978-825-16495

; Sequence 16495, Application US/09978825

; GENERAL INFORMATION:

; APPLICANT: Mitcham, Jennifer

; APPLICANT: Skeiky, Yasir

; APPLICANT: Persing, David

; APPLICANT: Bhatia, Ajay

; APPLICANT: Maisonneuve, Jean Francois

; APPLICANT: Zhang, Yanni

; APPLICANT: Wang, Siqing

; APPLICANT: Jen, Shiyun

; APPLICANT: Lodes, Michael

; APPLICANT: Benson, Darin

; APPLICANT: Jones, Robert

; APPLICANT: Carter, Darlick

; APPLICANT: Barth, Brenda

; APPLICANT: Douglass, John

; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne

; FILE REFERENCE: 210121.514C1

; CURRENT APPLICATION NUMBER: US/09/978, 825

; CURRENT FILING DATE: 2003-01-29

; NUMBER OF SEQ ID NOS: 30992

; SEQ ID NO 16495

; LENGTH: 59

; TYPE: PRT

; ORGANISM: Propionl acnes

US-09-978-825-16495

Query Match 100.0%; Score 27; DB 5; Length 59;

Best Local Similarity 100.0%; Pred. No. 1.2e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGAVP 5

Db 53 GGAVP 57

RESULT 9

US-10-057-498-16495

; Sequence 16495, Application US/10057498

; GENERAL INFORMATION:

; APPLICANT: Mitcham, Jennifer

; APPLICANT: Skeiky, Yasir

; APPLICANT: Persing, David

; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne

; FILE REFERENCE: 210121.514

; CURRENT APPLICATION NUMBER: US/10/057, 498

; CURRENT FILING DATE: 2001-04-20

; NUMBER OF SEQ ID NOS: 29212

; SEQ ID NO 16495

; LENGTH: 59

; TYPE: PRT

; ORGANISM: Propionl acnes

US-10-057-498-16495

Query Match 100.0%; Score 27; DB 6; Length 59;

Best Local Similarity 100.0%; Pred. No. 1.2e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGAVP 5

Db 53 GGAVP 57

RESULT 10

US-09-855-754B-11

; Sequence 11, Application US/09855754B

; GENERAL INFORMATION:

; APPLICANT: BOURSIAUX-EUDE, CAROLINE

; APPLICANT: GUISSO-MACLOUF, NICOLE

; TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED

; TITLE OF INVENTION: REGIONS OF PERTACTIN IN BORDETELLA PERTUSSIS,

; TITLE OF INVENTION: BORDETELLA PARAPERTUSSIS, AND BORDETELLA

; TITLE OF INVENTION: BRONCHISEPTICA, THEIR USE IN DIAGNOSTICS, AND IN

; FILE REFERENCE: 03495-0206-00000

; CURRENT APPLICATION NUMBER: US/09/855, 754B

; CURRENT FILING DATE: 2001-05-16

; PRIOR APPLICATION NUMBER: 60/206, 969

; NUMBER OF SEQ ID NOS: 25

; SOFTWARE: Patentin Ver. 2.1

; SEQ ID NO 11

; LENGTH: 61

; TYPE: PRT

; ORGANISM: Bordetella bronchiseptica

US-09-855-754B-11

Query Match 100.0%; Score 27; DB 5; Length 61;

Best Local Similarity 100.0%; Pred. No. 1.3e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGAVP 5

Db 13 GGAVP 17

RESULT 11

PCT-US02-32727-19550

; Sequence 19550, Application PC/TUS0232727

; GENERAL INFORMATION:

; APPLICANT: Mitcham, Jennifer

; APPLICANT: Skeiky, Yasir

; APPLICANT: Persing, David

PCT-US02-32727-19550

Query Match 100.0%; Score 27; DB 5; Length 61;

Best Local Similarity 100.0%; Pred. No. 1.3e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

; APPLICANT: Bhatia, Ajay
; APPLICANT: Maisonneuve, Jean Francois
; APPLICANT: Zhang, Yanni
; APPLICANT: Wang, Shiqing
; APPLICANT: Jen, Shyian
; APPLICANT: Lodes, Michael
; APPLICANT: Benson, Darin
; APPLICANT: Jones, Robert
; APPLICANT: Carter, Darrick
; APPLICANT: Barth, Brenda
; APPLICANT: Douglass, John
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne
; FILE REFERENCE: 210121.514C1
; CURRENT APPLICATION NUMBER: PCT/US02/32727
; CURRENT FILING DATE: 2002-10-11
; NUMBER OF SEQ ID NOS: 30992
; SEQ ID NO 19550
; LENGTH: 84
; TYPE: PRT
; ORGANISM: Propionibacterium
PCT-US02-32727-19550

```

Query Match

Best Local Similarity 100.0%; Score 27; DB 1; Length 84;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGAVP 5
DB 49 GGAVP 53

RESULT 12

```

; Sequence 978-825-19550
; Application US/09978825
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Skelky, Yasir
; APPLICANT: Persing, David
; APPLICANT: Bhatia, Ajay
; APPLICANT: Maisonneuve, Jean Francois
; APPLICANT: Zhang, Yanni
; APPLICANT: Wang, Shiqing
; APPLICANT: Jen, Shyian
; APPLICANT: Lodes, Michael
; APPLICANT: Benson, Darin
; APPLICANT: Jones, Robert
; APPLICANT: Carter, Darrick
; APPLICANT: Barth, Brenda
; APPLICANT: Douglass, John
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne
; FILE REFERENCE: 210121.514C1
; CURRENT APPLICATION NUMBER: US/09/978,825
; CURRENT FILING DATE: 2003-01-29
; NUMBER OF SEQ ID NOS: 30992
; SEQ ID NO 19550
; LENGTH: 84
; TYPE: PRT
; ORGANISM: Propionibacterium
US-09-978-825-19550

```

Query Match

Best Local Similarity 100.0%; Score 27; DB 5; Length 84;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGAVP 5
DB 49 GGAVP 53

RESULT 13

```

; Sequence 10-057-498-19550
; Application US/10057498
; GENERAL INFORMATION:

```

```

; APPLICANT: Mitcham, Jennifer
; APPLICANT: Skelky, Yasir
; APPLICANT: Persing, David
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne
; FILE REFERENCE: 210121.514
; CURRENT APPLICATION NUMBER: US/10/057,498
; CURRENT FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 29212
; SEQ ID NO 19550
; LENGTH: 84
; TYPE: PRT
; ORGANISM: Propionibacterium
US-10-057-498-19550

```

Query Match

Best Local Similarity 100.0%; Score 27; DB 6; Length 84;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGAVP 5
DB 49 GGAVP 53

RESULT 14

```

; Sequence 27641-27641
; Application PC/TUS0232727
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Skelky, Yasir
; APPLICANT: Persing, David
; APPLICANT: Bhatia, Ajay
; APPLICANT: Maisonneuve, Jean Francois
; APPLICANT: Zhang, Yanni
; APPLICANT: Wang, Shiqing
; APPLICANT: Jen, Shyian
; APPLICANT: Lodes, Michael
; APPLICANT: Benson, Darin
; APPLICANT: Jones, Robert
; APPLICANT: Carter, Darrick
; APPLICANT: Barth, Brenda
; APPLICANT: Douglass, John
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne
; FILE REFERENCE: 210121.514C1
; CURRENT APPLICATION NUMBER: PCT/US02/32727
; CURRENT FILING DATE: 2002-10-11
; NUMBER OF SEQ ID NOS: 30992
; SEQ ID NO 27641
; LENGTH: 105
; TYPE: PRT
; ORGANISM: Propionibacterium
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (83)
; OTHER INFORMATION: Xaa = Any Amino Acid
PCT-US02-32727-27641

```

Query Match

Best Local Similarity 100.0%; Score 27; DB 1; Length 105;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGAVP 5
DB 94 GGAVP 98

RESULT 15

```

; Sequence 27641-27641
; Application US/09978825
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Skelky, Yasir
; APPLICANT: Persing, David
; APPLICANT: Bhatia, Ajay

```

```

: APPLICANT: Maisonneuve, Jean Francois
: APPLICANT: Zhang, Yanni
: APPLICANT: Wang, Siqing
: APPLICANT: Jen, Shyian
: APPLICANT: Lodes, Michael
: APPLICANT: Benson, Darin
: APPLICANT: Jones, Robert
: APPLICANT: Carter, Darick
: APPLICANT: Barth, Brenda
: APPLICANT: Douglass, John
: TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acnes
: FILE REFERENCE: 210121.514C1
: CURRENT APPLICATION NUMBER: US/09/978,825
: CURRENT FILING DATE: 2003-01-29
: NUMBER OF SEQ ID NOS: 30992
: SEQ ID NO 27641
: LENGTH: 105
: TYPE: PRT
: ORGANISM: Propionibacterium
: FEATURE:
: NAME/KEY: unsure
: LOCATION: (83)
: OTHER INFORMATION: Xaa - Any Amino Acid
US-09-978-825-27641

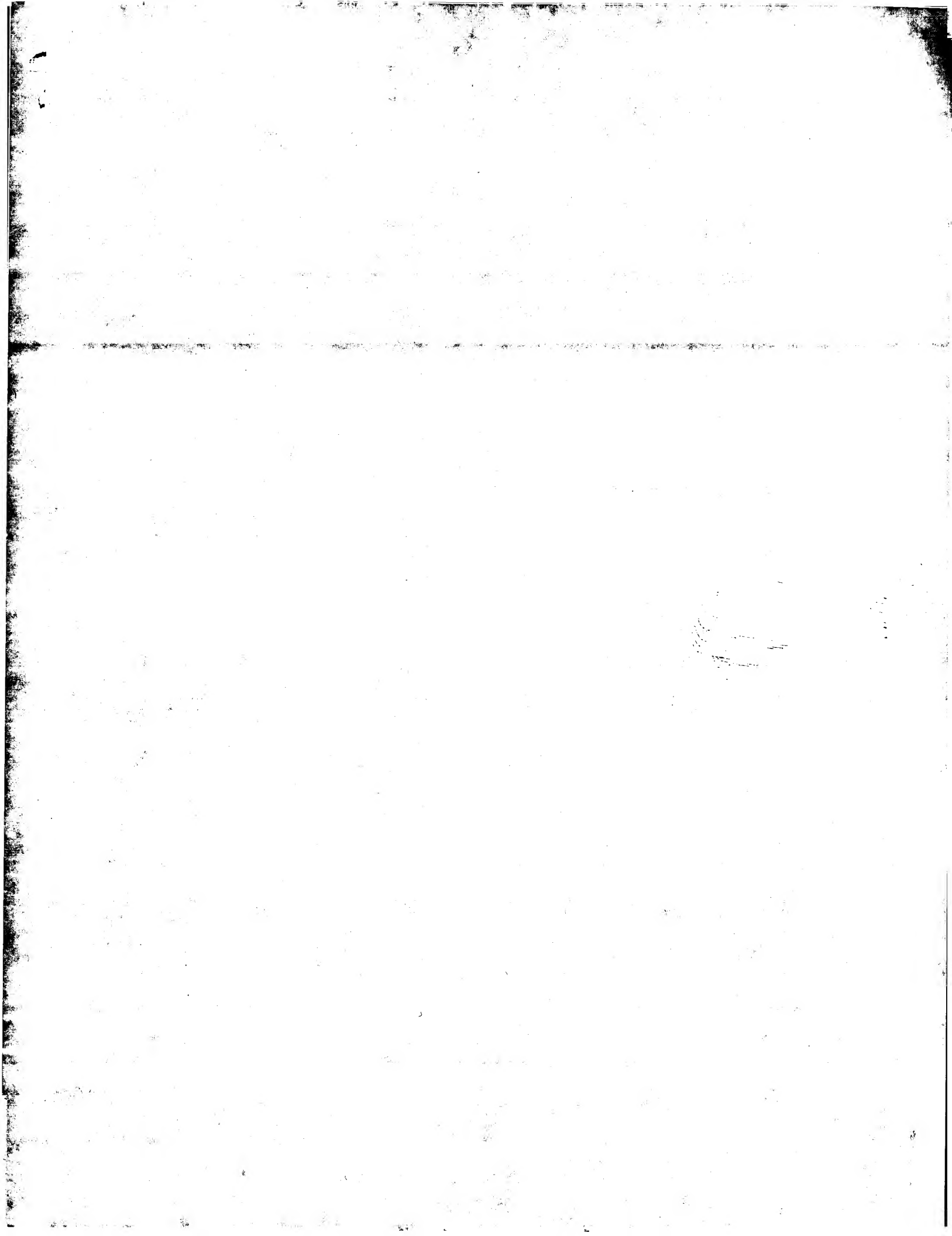
```

```

Query Match          100.0%; Score 27; DB 5; Length 105;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Caps 0;
QY      1 GGAVP 5
      11111
Db      94 GGAVP 98

```

Search completed: March 17, 2003, 12:20:54
Job time : 24 secs




```

XX  WO9419372-A.
XX
XX  01-SEP-1994.
XX
XX  08-FEB-1994; 94WO-US01393.
XX
XX  17-FEB-1993; 93US-0018604.
XX
XX  (SCIO-) SCIOS NOVA INC.
XX
XX  Chakravarty S, Kyle DJ;
XX
XX  WPI: 1994-294261/36.
XX
XX  New cyclic bradykinin antagonist peptide(s) - used for treating
XX  e.g. pain, inflammation, angina, arthritis, asthma, allergies,
XX  rhinitis or shock
XX
XX  Claim 5; Page 41; 47pp; English.
XX
XX  New bradykinin analogues are disclosed which include intramolecular
XX  cyclisation via side chains to provide a rigid structure having
XX  enhanced affinity for the bradykinin receptors. Specifically, amino
XX  acid substitutions are made at positions 2 and 5 or 6 to facilitate
XX  cyclisation of the peptide through covalent bonding. Examples of
XX  suitable substitutions are Cys for disulphide bonding and Lys with
XX  Glu or Asp for amide bond formation. An additional residue is added
XX  to the N-terminal, and additional substitutions may be made at other
XX  positions.
XX  The peptides are potent bradykinin receptor antagonists which may
XX  be used for treating local pain and inflammation from burns, wounds,
XX  cuts, rashes, etc., and also for treating angina, arthritis,
XX  asthma, allergies, rhinitis, shock, inflammatory bowel disease and
XX  low blood pressure.
XX  The present sequence is a specifically claimed example of the
XX  disclosed peptides
XX
XX  Sequence 10 AA;
XX
SQ
XX
XX  Query Match
XX  Best Local Similarity 100.0%; Score 31; DB 15; Length 10;
XX  Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 GGFDP 5
XX  |||||
XX  4 GGFDP 8
XX
DB
XX
XX  RESULT 2
XX  ID ABA47180 standard; peptide; 11 AA.
XX
XX  ABA47180;
XX
XX  19-AUG-2002 (first entry)
XX
XX  Human Blys binding scfv VH CDR3 SEQ ID 3191.
XX
XX  Blys; B lymphocyte stimulator; TNF superfamily; human; cytostatic;
XX  tumour necrosis factor; B cell proliferation; B cell differentiation;
XX  immunosuppressive; immunostimulant; immunomodulatory; antipneumatic;
XX  antiAIDS; vaccine; cancer; immune; autoimmune disorder; immunodeficiency;
XX  systemic lupus erythematosus; rheumatoid arthritis; CVID; AIDS;
XX  common variable immunodeficiency; acquired immunodeficiency syndrome.
XX
XX  Homo sapiens.
XX
XX  WO200202641-A1.
XX
XX  10-JAN-2002.
XX
XX  15-JUN-2001; 2001WO-US19110.
XX

```

```

XX  16-JUN-2000; 2000US-212210P.
XX
XX  17-OCT-2000; 2000US-240816P.
XX
XX  16-MAR-2001; 2001US-276248P.
XX
XX  21-MAR-2001; 2001US-277379P.
XX
XX  25-MAY-2001; 2001US-293499P.
XX
XX  (HUMA-) HUMAN GENOME SCI INC.
XX  (CAMP-) CAMBRIDGE ANTIBODY TECHNOLOGY.
XX
XX  Ruben SM, Barash SC, Choi GH, Vaughan T, Hilbert D;
XX
XX  WPI: 2002-114799/15.
XX
XX  Antibodies against B lymphocyte stimulating polypeptides, useful for
XX  the diagnosis and treatment of cancers and immune disorders -
XX
XX  Claim 2; Page 3131; 3148pp; English.
XX
XX  This invention describes novel antibodies that immunospecifically bind to
XX  B lymphocyte stimulator (Blys) polypeptides. Blys is a member of the
XX  tumour necrosis factor (TNF) super family and induces B cell
XX  proliferation and differentiation. The antibodies of the invention have
XX  cytostatic, immunosuppressive, immunostimulant, immunomodulatory,
XX  antipneumatic and antiAIDS activity and can be used in vaccines to
XX  inhibit the expression and activity of Blys. The antibodies bind to Blys
XX  and so may be used to detect and quantify the presence of Blys in
XX  biological samples and may be used in this way to diagnose disease
XX  associated with aberrant expression of Blys. They may also be
XX  administered to treat diseases associated with aberrant Blys expression
XX  and activity such as cancer, immune, and autoimmune disorders and
XX  diseases, e.g. systemic lupus erythematosus, rheumatoid arthritis,
XX  immunodeficiency (e.g. common variable immunodeficiency (CVID) and
XX  acquired immunodeficiency syndrome (AIDS)). ABA43990-ABA47228 represent
XX  the antibodies and fragments of the antibodies described in the method
XX  of the invention.
XX
XX  Sequence 11 AA;
XX
SQ
XX
XX  Query Match
XX  Best Local Similarity 100.0%; Score 31; DB 23; Length 11;
XX  Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 GGFDP 5
XX  |||||
XX  7 GGFDP 11
XX
DB
XX
XX  RESULT 3
XX  ID AAE16188 standard; peptide; 56 AA.
XX
XX  AAE16188;
XX
XX  26-MAR-2002 (first entry)
XX
XX  B. bronchiseptica strain I-3 pertactin outer membrane protein region I.
XX
XX  Pertactin; PRN; outer membrane protein; vaccine; Bordetella infection;
XX  therapy; antiheliotic; antibacterial; region I.
XX
XX  Bordetella bronchiseptica.
XX
XX  WO200190143-A2.
XX
XX  29-NOV-2001.
XX
XX  23-MAY-2001; 2001WO-EP06457.
XX
XX  25-MAY-2000; 2000US-206969P.
XX
XX  (INSP ) INST PASTEUR.
XX

```

PI Guiso-maclouf N, Boursaux-eude C;
 XX WPI; 2002-097639/13.
 XX Polypeptides containing polymorphisms of the repeated regions of
 PT pertactin in Bordetella species, useful in immunogenic compositions for
 PT treating infections caused by Bordetella and in diagnostic methods
 XX
 PS Claim 26; Fig 1b; 47pp; English.
 XX
 CC The present invention relates to Bordetella bronchiseptica pertactin
 CC (outer membrane protein) or their fragments. Pertactin (PRN) is used as
 CC vaccine. Pertactin antibody is useful for treating Bordetella infections
 CC and used to detect Bordetella antigens in biological preparations or in
 CC purifying corresponding proteins, glycoproteins or their mixtures when
 CC used in affinity chromatographic columns. Pertactin is useful as antigens
 CC to identify antibodies to Bordetella in materials such as human or other
 CC animal tissue and human or other animal cells, as well as biological
 CC fluids, such as human or other animal body fluids, including human sera,
 CC and to determine the concentration of Ab in those materials. Thus the
 CC antigens can be used for qualitative or quantitative determination of
 CC Bordetella in a material. The present sequence is B. bronchiseptica
 CC pertactin outer membrane protein region I.
 CC
 SQ Sequence 56 AA;

Query Match 100.0%; Score 31; DB 23; Length 56;
 Best Local Similarity 100.0%; Pred. No. 44;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGDP 5
 |||||
 DB 23 GGDP 27

RESULT 4
 AAY76003
 ID AAY76003 standard; Protein; 60 AA.
 XX
 AC AAY76003;

DE 27-MAR-2000 (first entry)
 XX
 DE Murine skin cell protein, SEQ ID NO:181.

XX Skin; dermal papilla; keratinocyte; neonatal foreskin fibroblast;
 KM embryonic skin cell; keratinocyte stem cell; transit amplifying cell;
 KM secreted; transmembrane; inflammation; cancer; neurological disease;
 KM angiogenesis; tumour vascularisation; growth disorder;
 KM developmental disorder; skin wound; hair follicle disorder;
 KM anti-inflammatory; cytostatic; neuroprotective; vulnery.
 XX
 OS Mus sp.

XX WO9955865-A1.

PD 04-NOV-1999.

PF 29-APR-1999; 99MO-NZ00051.

PR 29-APR-1998; 98US-0069726.
 PR 09-NOV-1998; 98US-0188930.

PA (GENE-) GENESIS RES & DEV CORP LTD.

PI Strachan L, Sleeman M, Watson JD, Onrust R, Kumble A, Murison JG;
 XX WPI; 2000-072177/06.

PT Novel polynucleotides useful for the treatment of various conditions
 PT including wounds and cancer -
 XX
 PS Claim 4; Page 124; 235pp; English.

XX The invention relates to novel nucleic acid sequences derived from rat
 CC dermal papilla, human keratinocytes and neonatal foreskin fibroblasts,
 CC and mouse embryonic skin, keratinocyte stem cells and transit amplifying
 CC cells. Polypeptides of the invention may be used to treat inflammation,
 CC cancer and neurological diseases. The proteins may be used to stimulate
 CC the growth and motility of keratinocytes, to inhibit the growth of
 CC cancer cells, to modulate angiogenesis and tumour vascularisation, to
 CC inhibit binding of HIV-1 to leukocytes. The invention may also be used
 CC to treat growth and developmental defects, skin wounds and hair follicle
 CC disorders. Sequences AAY75942-Y76123 represent polypeptides encoded
 CC by cDNA sequences derived from several mouse, rat or human skin cell
 CC types. Sequences AAY75942-Y75947, AAY76020-Y76021, AAY76094-Y76104 and
 CC AAY76119 are proteins with an N-terminal signal sequence, indicating
 CC that they are secreted. Sequences AAY75986-Y75989, AAY76061-Y76071,
 CC AAY76106-Y76109 and AAY76121-Y76122 are proteins with one or more
 CC putative transmembrane domains.
 XX

Query Match 100.0%; Score 31; DB 21; Length 60;
 Best Local Similarity 100.0%; Pred. No. 47;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGDP 5
 |||||
 DB 51 GGDP 55

RESULT 5
 AAY76065
 ID AAY76065 standard; Protein; 60 AA.
 XX
 AC AAY76065;

DE 27-MAR-2000 (first entry)

DE Murine skin cell transmembrane protein, SEQ ID NO:320.

XX Skin; dermal papilla; keratinocyte; neonatal foreskin fibroblast;
 KM embryonic skin cell; keratinocyte stem cell; transit amplifying cell;
 KM secreted; transmembrane; inflammation; cancer; neurological disease;
 KM angiogenesis; tumour vascularisation; growth disorder;
 KM developmental disorder; skin wound; hair follicle disorder;
 KM anti-inflammatory; cytostatic; neuroprotective; vulnery.
 XX
 OS Mus sp.

XX WO9955865-A1.

PD 04-NOV-1999.

PF 29-APR-1999; 99MO-NZ00051.

PR 29-APR-1998; 98US-0069726.
 PR 09-NOV-1998; 98US-0188930.

PA (GENE-) GENESIS RES & DEV CORP LTD.

PI Strachan L, Sleeman M, Watson JD, Onrust R, Kumble A, Murison JG;
 XX WPI; 2000-072177/06.

PT Novel polynucleotides useful for the treatment of various conditions
 PT including wounds and cancer -
 XX
 PS Claim 4; Page 188; 235pp; English.

XX The invention relates to novel nucleic acid sequences derived from rat
 CC dermal papilla, human keratinocytes and neonatal foreskin fibroblasts,
 CC and mouse embryonic skin, keratinocyte stem cells and transit amplifying
 CC cells. Polypeptides of the invention may be used to treat inflammation,

CC Cancer and neurological diseases. The proteins may be used to stimulate
 CC the growth and motility of keratinocytes, to inhibit the growth of
 CC cancer cells, to modulate angiogenesis and tumour vascularisation, to
 CC modulate skin inflammation, to modulate epithelial cell growth and to
 CC inhibit binding of HIV-1 to leukocytes. The invention may also be used
 CC to treat growth and developmental defects, skin wounds and hair follicle
 CC disorders. Sequences AAY75942-Y76123 represent polypeptides encoded
 CC by cDNA sequences derived from several mouse, rat or human skin cell
 CC types. Sequences AAY75942-Y75947, AAY76020-Y76021, AAY76094-Y76104 and
 CC AAY76119 are proteins with an N-terminal signal sequence, indicating
 CC that they are secreted. Sequences AAY75986-Y75989, AAY76061-Y76071,
 CC AAY76106-Y76109 and AAY76121-Y76122 are proteins with one or more
 CC putative transmembrane domains.

XX Sequence 60 AA;

Query Match 100.0%; Score 31; DB 21; Length 60;
 Best Local Similarity 100.0%; Pred. No. 47;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFDP 5
 |||||
 Db 51 GGFDP 55

RESULT 6
 AAB55942
 ID AAB55942 standard; Protein; 60 AA.

AC AAB55942;

DT 08-MAR-2001 (first entry)

DE Skin cell protein, SEQ ID NO: 181.

XX Mouse; skin cell; cytosolic; antiinflammatory; anti-HIV;
 KW nootropic; neuroprotective; vulnerrary; immunomodulatory; vaccine;
 KM keratinocyte growth stimulation; cancer; angiogenesis inhibition;
 XX inflammation; neurological disease.

OS Mus sp.

XX WO200069884-A2.

PD 23-NOV-2000.

PF 15-MAY-2000; 2000WO-NZ00075.

XX 14-MAY-1999; 99US-0312283.

PA (GENE-) GENESIS RES & DEV CORP LTD.

PI Watson JD, Strachan L, Orrust R, Sleeman M, Kumble KD, Murison JG;

DR WPI: 2001-007495/01.

DR N-PSDB; AAC99600.

XX New isolated polynucleotide used in the identification of genetic
 PT disorders and encoding polypeptides used for treating inflammatory
 PT disease, cancer and neurological diseases -

XX Claim 4; Page 161; 352pp; English.

XX The present sequence is a polypeptide which is expressed in
 CC mammalian skin cells. The polypeptide is useful for stimulating
 CC keratinocyte growth and motility, inhibiting the growth of cancer cells,
 CC modulating angiogenesis, inhibiting angiogenesis and vascularisation of
 CC tumours, modulating skin inflammation, stimulating the growth of
 CC epithelial cells, inhibiting the binding of human immunodeficiency virus
 CC (HIV)-1 to leukocytes, and treating inflammatory disease, cancer and
 CC neurological diseases. The polynucleotide can be used as a marker, in
 CC the identification of genetic disorders, and for the design of
 CC oligonucleotides for examining expression patterns.

XX Sequence 60 AA;

Query Match 100.0%; Score 31; DB 22; Length 60;
 Best Local Similarity 100.0%; Pred. No. 47;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFDP 5
 |||||
 Db 51 GGFDP 55

RESULT 7

AAB56004
 ID AAB56004 standard; Protein; 60 AA.

AC AAB56004;

DT 08-MAR-2001 (first entry)

DE Skin cell protein, SEQ ID NO: 320.

XX Mouse; skin cell; cytosolic; antiinflammatory; anti-HIV;
 KW nootropic; neuroprotective; vulnerrary; immunomodulatory; vaccine;
 KM keratinocyte growth stimulation; cancer; angiogenesis inhibition;
 XX inflammation; neurological disease.

OS Mus sp.

XX WO200069884-A2.

PD 23-NOV-2000.

PF 15-MAY-2000; 2000WO-NZ00075.

XX 14-MAY-1999; 99US-0312283.

PA (GENE-) GENESIS RES & DEV CORP LTD.

PI Watson JD, Strachan L, Orrust R, Sleeman M, Kumble KD, Murison JG;

DR WPI: 2001-007495/01.

DR N-PSDB; AAC99703.

XX New isolated polynucleotide used in the identification of genetic
 PT disorders and encoding polypeptides used for treating inflammatory
 PT disease, cancer and neurological diseases -

XX Claim 4; Page 254; 352pp; English.

XX The present sequence is a polypeptide which is expressed in
 CC mammalian skin cells. The polypeptide is useful for stimulating
 CC keratinocyte growth and motility, inhibiting the growth of cancer cells,
 CC modulating angiogenesis, inhibiting angiogenesis and vascularisation of
 CC tumours, modulating skin inflammation, stimulating the growth of
 CC epithelial cells, inhibiting the binding of human immunodeficiency virus
 CC (HIV)-1 to leukocytes, and treating inflammatory disease, cancer and
 CC neurological diseases. The polynucleotide can be used as a marker, in
 CC the identification of genetic disorders, and for the design of
 CC oligonucleotides for examining expression patterns.

XX Sequence 60 AA;

Query Match 100.0%; Score 31; DB 22; Length 60;
 Best Local Similarity 100.0%; Pred. No. 47;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFDP 5
 |||||
 Db 51 GGFDP 55

RESULT 8

ABR72142
ID ABR72142 standard; Protein: 60 AA.
XX
AC ABR72142;
XX
DT 04-APR-2002 (first entry)
XX
DE Murine protein isolated from skin cells SEQ ID NO: 181.
XX
KW Human; rat; mouse; skin cell; skin wound; cancer; growth defect;
KW developmental defect; inflammatory disease; dermatological; vulnary;
KW immunomodulator; anti-inflammatory; cytostatic; neuroprotective.
XX
OS Mus sp.
XX
PN WO200190357-A1.
XX
PD 29-NOV-2001.
XX
PF 24-MAY-2001; 2001WO-NZ00099.
XX
PR 24-MAY-2000; 2000US-206650P.
PR 25-JUL-2000; 2000US-221232P.
XX
PA (GENE-) GENESIS RES & DEV CORP LTD.
XX
PI Watson JD, Strachan L, Sleeman M, Onrust R, Murison JG, Kumble KD;
XX
DR WPI: 2002-122020/16.
XX
PT New polynucleotides and polypeptides encoded by the polynucleotides
PT isolated from skin cells, useful for treating skin wounds, cancers,
PT growth and developmental defects, inflammatory diseases, or for
PT modulating immune responses -
XX
PS Example 2; Page 142; 466pp; English.
XX
CC The present invention provides the protein and coding sequences of cDNAs
CC isolated from human, murine and rat skin cell libraries. The sequences
CC can be used in the development of therapeutic agents useful in the
CC treatment of skin diseases, including skin wounds, cancer, growth
CC defects, developmental defects and inflammatory diseases. The proteins
CC have important roles in the induction of hair growth, cell proliferation
CC and cell-cell interaction, in maintaining tissue integrity, in wound
CC healing and in modulating immune responses. The present sequence is a
CC polypeptide of the invention.
XX
SQ Sequence 60 AA:
XX
Query Match 100.0%; Score 31; DB 23; Length 60;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 GCFDP 5
Db 51 GCFDP 55
XX
RESULT 9
ABR72204
ID ABR72204 standard; Protein: 60 AA.
XX
AC ABR72204;
XX
DT 04-APR-2002 (first entry)
XX
DE Murine protein isolated from skin cells SEQ ID NO: 320.
XX
KW Human; rat; mouse; skin cell; skin wound; cancer; growth defect;
KW developmental defect; inflammatory disease; dermatological; vulnary;
KW immunomodulator; anti-inflammatory; cytostatic; neuroprotective.
XX
OS Mus sp.

XX
PN WO200190357-A1.
XX
PD 29-NOV-2001.
XX
PF 24-MAY-2001; 2001WO-NZ00099.
XX
PR 24-MAY-2000; 2000US-206650P.
PR 25-JUL-2000; 2000US-221232P.
XX
PA (GENE-) GENESIS RES & DEV CORP LTD.
XX
PI Watson JD, Strachan L, Sleeman M, Onrust R, Murison JG, Kumble KD;
XX
DR WPI: 2002-122020/16.
XX
PT New polynucleotides and polypeptides encoded by the polynucleotides
PT isolated from skin cells, useful for treating skin wounds, cancers,
PT growth and developmental defects, inflammatory diseases, or for
PT modulating immune responses -
XX
PS Example 2; Page 210; 466pp; English.
XX
CC The present invention provides the protein and coding sequences of cDNAs
CC isolated from human, murine and rat skin cell libraries. The sequences
CC can be used in the development of therapeutic agents useful in the
CC treatment of skin diseases, including skin wounds, cancer, growth
CC defects, developmental defects and inflammatory diseases. The proteins
CC have important roles in the induction of hair growth, cell proliferation
CC and cell-cell interaction, in maintaining tissue integrity, in wound
CC healing and in modulating immune responses. The present sequence is a
CC polypeptide of the invention.
XX
SQ Sequence 60 AA:
XX
Query Match 100.0%; Score 31; DB 23; Length 60;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 GCFDP 5
Db 51 GCFDP 55
XX
RESULT 10
AAB51430
ID AAB51430 standard; Protein: 83 AA.
XX
AC AAB51430;
XX
DT 16-FEB-2001 (first entry)
XX
DE Human secreted protein BLAST search protein SEQ ID NO: 107.
XX
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; anticancer;
KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein.
XX
OS Homo sapiens.
XX
PN WO200058495-A1.
XX
PD 05-OCT-2000.
XX
PF 23-MAR-2000; 2000WO-US07661.
XX
PR 26-MAR-1999; 99US-0126504.
PR 07-JAN-2000; 2000US-0174847.
XX
PA (HDMA-) HUMAN GENOME SCI INC.
XX

PI Rosen CA, Ruben SM, Komatsoulis G;
 XX WPI; 2000-611720/58.
 XX
 PT New nucleic acid molecules encoding 45 human secreted proteins for
 PT diagnosing, preventing, treating or ameliorating medical conditions and
 XX used as food additives or preservatives -
 XX
 PS Disclosure; Page 374; 410pp; English.

CC The invention relates to the isolation of genes AAC93310-C93354 encoding
 CC the human secreted proteins AAB51380-B51423. The genes and proteins are
 CC useful for preventing, ameliorating or treating medical conditions, e.g.
 CC by protein or gene therapy. The genes are isolated from a range of human
 CC tissues disclosed in the specification. The nucleic acids, proteins,
 CC antibodies and (ant)agonists are useful in the diagnosis, treatment and
 CC prevention of: (a) cancer, e.g. breast and ovarian cancer, and other
 CC cancers of the adrenal gland, bone, bone marrow, breast,
 CC gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders
 CC e.g. Addison's disease, allergies, autoimmune haemolytic anaemia,
 CC autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple
 CC sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
 CC cardiovascular disorders such as myocardial ischaemias; (d) wound
 CC healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;
 CC and (f) infectious diseases such as viral, bacterial, fungal and
 CC parasitic infections. The present sequence is a protein isolated in the
 CC present invention.

Sequence 83 AA;

Query Match 100.0%; Score 31; DB 21; Length 83;
 Best Local Similarity 100.0%; Pred. NO. 64;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFDP 5
 |||||
 Db 51 GGFDP 55

RESULT 11
 AAB51431
 ID AAB51431 standard; Protein; 83 AA.

AC AAB51431;
 XX
 DT 16-FEB-2001 (first entry)

DE Human secreted protein BLAST search protein SEQ ID NO: 108.

KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW anti-allergic; hepatotropic; antidiabetic; anti-inflammatory; antitumor;
 KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein.

OS Homo sapiens.

XX
 PN WO200058495-A1.

PD 05-OCT-2000.

PF 23-MAR-2000; 2000WO-US07661.

PR 26-MAR-1999; 99US-0126504.

PR 07-JAN-2000; 2000US-0174847.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Ruben SM, Komatsoulis G;

WPI; 2000-611720/58.

PT New nucleic acid molecules encoding 45 human secreted proteins for

PT diagnosing, preventing, treating or ameliorating medical conditions and
 PT used as food additives or preservatives -
 XX
 PS Disclosure; Pages 374-375; 410pp; English.

CC The invention relates to the isolation of genes AAC93310-C93354 encoding
 CC the human secreted proteins AAB51380-B51423. The genes and proteins are
 CC useful for preventing, ameliorating or treating medical conditions, e.g.
 CC by protein or gene therapy. The genes are isolated from a range of human
 CC tissues disclosed in the specification. The nucleic acids, proteins,
 CC antibodies and (ant)agonists are useful in the diagnosis, treatment and
 CC prevention of: (a) cancer, e.g. breast and ovarian cancer, and other
 CC cancers of the adrenal gland, bone, bone marrow, breast,
 CC gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders
 CC e.g. Addison's disease, allergies, autoimmune haemolytic anaemia,
 CC autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple
 CC sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
 CC cardiovascular disorders such as myocardial ischaemias; (d) wound
 CC healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;
 CC and (f) infectious diseases such as viral, bacterial, fungal and
 CC parasitic infections. The present sequence is a protein isolated in the
 CC present invention.

Sequence 83 AA;

Query Match 100.0%; Score 31; DB 21; Length 83;
 Best Local Similarity 100.0%; Pred. NO. 64;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFDP 5
 |||||
 Db 51 GGFDP 55

RESULT 12
 AAB51438
 ID AAB51438 standard; Protein; 83 AA.

AC AAB51438;
 XX
 DT 16-FEB-2001 (first entry)

DE Human secreted protein BLAST search protein SEQ ID NO: 115.

KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW anti-allergic; hepatotropic; antidiabetic; anti-inflammatory; antitumor;
 KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein.

OS Homo sapiens.

XX
 PN WO200058495-A1.

PD 05-OCT-2000.

PF 23-MAR-2000; 2000WO-US07661.

PR 26-MAR-1999; 99US-0126504.

PR 07-JAN-2000; 2000US-0174847.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Ruben SM, Komatsoulis G;

WPI; 2000-611720/58.

PT New nucleic acid molecules encoding 45 human secreted proteins for
 PT diagnosing, preventing, treating or ameliorating medical conditions and
 XX used as food additives or preservatives -
 XX
 PS Disclosure; Page 377; 410pp; English.

Sequence 83 AA;

QY	1	GGEDP	5
	11111		
Db	51	GGEDP	55

RESULT 13
AAB51439
ID AAB51439 standard; Protein; 83 AA

DT 16-FEB-2001 (first entry)

Human secreted protein BLAST search protein SEQ ID NO: 116.

KM cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KM antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
KM vulnereary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KM cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder
KM neurological disease; infection; human; secreted protein.

OS Homo sapiens.

PN W0200058495-A1.

PD 05-OCT-2000.

PF 23-MAR-2000; 2000WO-US07661.

PR 26-MAR-1999; 99US-0126504.

XX XX

2 XX
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841

DR WPI; 2000-611720/58.

PT New nucleic acid molecules encoding 45 human secreted proteins for
PT diagnosing, preventing, treating or ameliorating medical conditions and
PT used as food additives or preservatives -

PS Disclosure; Pages 377-378; 410pp; English.

CC The invention relates to the isolation of genes AAC93310-C93354 encoding
CC the human secreted proteins AAB51380-B51423. This sequence represents a
CC fragment of the protein encoded by the gene isolated in the present
CC invention. The sequence is used as a query sequence for doing BLASTX

searches to determine homologous sequence to the protein. The genes and proteins are useful for preventing, ameliorating or treating medical conditions, e.g. by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification. The nucleic acids, proteins, antibodies and (ant)agonists are useful in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer, and other cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders such as myocardial ischaemias; (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such as viral, bacterial, fungal and parasitic infections.

SQ Sequence 83 AA;

Query Match	100.0%	Score 31;	DB 21;	Length 83;
Best Local Similarity	100.0%	Pred. No. 64;		
Matches	5;	Conservative	0;	Indels 0;
		Mismatches	0;	Gaps 0;

Qy	1	GGRDP	5
	1	1	1
	1	1	1
Db	51	GGRDP	55

RESULT 14
ABP40272
ID ABP40272 standard; Protein; 91 AA

DT 24-JUL-2002 (first entry)

DE Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:5117

KN Staphylococcus epidermidis; open reading frame; ORF; bacterial infection,
antibacterial; gene therapy.

OS Staphylococcus epidermidis.

US6380370-B1

PD 30-APR-2002.

PF 13-AUG-1998; 98US-0134001.

PR 14-AUG-1997; 97US-055779P.

PA (GENO-) GENOME THERAPEUTICS CORP

PI Doucette-Stamm LA, Bush D;

DR WPI; 2002-381255/41.

[illegible]

PT Novel isolated nucleic acid encoding a Staphylococcus epidermidis polypeptide, useful for diagnosing and treating bacterial infections -
XX
PS Disclosure: SEQ ID 5117; 267pp; English.

CC ABN90538 to ABN93374 represent *Staphylococcus aureus* strains.

CC ABN90538 o ABN933374 represent *Staphylococcus epidermidis* open reading
CC frame (ORF) nucleic acid sequences which encode the amino acid sequences
CC given in ABP35124 to ABP37960. The *S. epidermidis* sequences have
CC antibacterial activity and can be used in gene therapy. The sequences
CC can also be used in the diagnosis and treatment of bacterial infections,
CC particularly *S. epidermidis* infections. The sequences can be used to
CC screen for compounds able to interfere with the *S. epidermidis* life
CC cycle or inhibit *S. epidermidis* infection.
CC M.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the

CC USPTO web site.

XX Sequence 91 AA;

Query Match 100.0%; Score 31; DB 23; Length 91;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFPD 5

Db 72 GGFPD 76

RESULT 15

AAW21829 ID AAW21829 standard; Protein; 95 AA.

AC AAW21829;

DT 26-OCT-1997 (first entry)

DE Human DQA1*0101 class II MHC alpha chain.

KW Immunosuppression; major histocompatibility complex class II;

KW MHC; mixed lymphocyte reaction; allorecognition; cytotoxic T cell;

KW alloimmunity; autoimmune disease; organ transplant;

KW multiple sclerosis; rheumatoid arthritis; DQA1*0101.

OS Homo sapiens.

Key Location/Qualifiers

Region 87..103 /note="alpha chain aa62-77 conserved region (Claim 11)"

WO9710711-A1.

PD 27-MAR-1997.

PF 23-SEP-1996; 96WO-US15662.

PR 21-SEP-1995; 95US-0004117.

PA (AUTO-) AUTOIMMUNE INC.

PI Carpenter CB, Murphy BT, Sayegh MH;

DR WPI; 1997-202534/18.

Suppression of immune responses with major histocompatibility complex class II peptide(s) - useful in alloimmunity; e.g. organ transplantation, and autoimmunity; e.g. in multiple sclerosis or rheumatoid arthritis

Claim 10; Fig 1A-B; 55pp; English.

The sequence comprises the first 95 amino acid residues of the alpha chain of the class II MHC from human DQA1*0101. It contains a highly conserved region that is maintained across alleles and species (see also AAW21827 and AAW21828). A novel method of suppressing an immune response comprises administering a class II MHC alpha chain, or a fragment (see also AAW21830-32), that can suppress at least one of the following: a mixed lymphocyte reaction or other T-cell allorecognition reaction; generation of cytotoxic T-cells recognising an alloantigen; lymphocyte proliferation against tissue antigen; and stimulatory cytokine production by lymphocytes. Immune responses that can be abated or suppressed include alloimmunity (e.g. in organ transplantation) and autoimmunity (e.g. in multiple sclerosis or rheumatoid arthritis). Immune responses can be down-regulated specifically without the adverse effects of conventional treatments.

Sequence 95 AA;

Query Match 100.0%; Score 31; DB 18; Length 95;
Best Local Similarity 100.0%; Pred. No. 73;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFPD 5

Db 55 GGFPD 59

Search completed: March 17, 2003, 12:08:55
Job time : 32.3333 secs

GenCore version 5.1.4-p5-4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 17, 2003, 12:06:59 ; Search time 12 Seconds
(without alignments)
12.260 Million cell updates/sec

Title: SEQ25-FD
Perfect score: 31
Sequence: 1 gfdp 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/pdata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/pdata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/pdata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/pdata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/pdata/2/1aa/CTUS.COMB.pep:*
6: /cgn2_6/pdata/2/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31	100.0	60	4	US-09-188-930-181
2	31	100.0	60	4	US-09-188-930-320
3	31	100.0	91	4	US-09-134-001C-5117
4	31	100.0	117	3	US-08-545-809A-110
5	31	100.0	255	2	US-08-484-905-110
6	31	100.0	255	3	US-08-481-985B-110
7	31	100.0	255	3	US-08-370-476-110
8	28	90.3	21	1	US-08-047-033-12
9	28	90.3	2475	4	US-09-413-814-48
10	27	87.1	116	2	US-08-846-134-1
11	27	87.1	317	6	5340934-11
12	27	87.1	466	4	US-09-470-512A-9
13	27	87.1	932	4	US-08-887-534A-45
14	27	87.1	1066	2	US-08-633-770A-1
15	27	87.1	1070	2	US-08-633-770A-2
16	26	83.9	28	1	US-08-124-369-3
17	26	83.9	132	1	US-08-647-960-4
18	26	83.9	167	1	US-08-470-892-5
19	26	83.9	167	1	US-08-485-069-5
20	26	83.9	232	1	US-08-425-763-2
21	26	83.9	232	3	US-07-934-373C-38
22	26	83.9	232	3	US-08-437-642B-38
23	26	83.9	232	3	US-08-811-757-2
24	26	83.9	232	3	US-09-249-230-2
25	26	83.9	232	5	PCT-US93-07832-38
26	26	83.9	326	4	US-09-333-611-12
27	26	83.9	454	2	US-07-934-373C-22

28	26	83.9	454	3	US-08-437-642B-22	Sequence 22, Appl
29	26	83.9	454	4	US-08-146-206C-22	Sequence 22, Appl
30	26	83.9	454	5	PCT-US93-07832-22	Sequence 22, Appl
31	26	83.9	459	6	5194375-6	Patent No. 5194375
32	26	83.9	689	4	US-08-778-570B-16	Sequence 16, Appl
33	26	83.9	689	4	US-09-059-584-16	Sequence 16, Appl
34	26	83.9	709	4	US-08-778-570B-15	Sequence 15, Appl
35	26	83.9	709	4	US-09-059-584-15	Sequence 15, Appl
36	26	83.9	3080	6	5223423-4	Patent No. 5223423
37	25	80.6	6	1	US-07-718-577-22	Sequence 22, Appl
38	25	80.6	7	6	5473052-24	Patent No. 5473052
39	25	80.6	9	3	US-08-159-339A-1012	Sequence 1012, Ap
40	25	80.6	11	1	US-08-107-684B-20	Sequence 20, Appl
41	25	80.6	11	1	US-08-107-684B-26	Sequence 26, Appl
42	25	80.6	12	4	US-09-222-779-1	Sequence 1, Appl1
43	25	80.6	13	4	US-09-222-779-2	Sequence 2, Appl1
44	25	80.6	19	3	US-08-895-707-17	Sequence 17, Appl
45	25	80.6	21	1	US-08-102-738-4	Sequence 4, Appl1

ALIGNMENTS

```

RESULT 1
US-09-188-930-181
; Sequence 181, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Orrust, Rene
; APPLICANT: Morrison, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE REFERENCE: 11000.1011c1
; CURRENT APPLICATION NUMBER: US/09/188,930A
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 181
; LENGTH: 60
; TYPE: PRT
; ORGANISM: mouse
US-09-188-930-320
; Sequence 320, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Orrust, Rene
; APPLICANT: Morrison, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE REFERENCE: 11000.1011c1
; CURRENT APPLICATION NUMBER: US/09/188,930A
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 320
; LENGTH: 60
; TYPE: PRT
; ORGANISM: mouse

Query Match      100.0%  Score 31:  DB 4:  Length 60:
Best Local Similarity 100.0%  Pred. No. 12:
Matches 5;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;
QY      1 GGFDP 5
DB      51 GGFDP 55

```

LENGTH: 60
TYPE: PRT
ORGANISM: Mouse
US-09-188-930-320

Query Match
Best Local Similarity 100.0%; Score 31; DB 4; Length 60;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFDP 5
|||||
DB 51 GGFDP 55

RESULT 3
US-09-134-001C-5117
Sequence 5117, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 5117
LENGTH: 91
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5117

Query Match
Best Local Similarity 100.0%; Score 31; DB 4; Length 91;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFDP 5
|||||
DB 72 GGFDP 76

RESULT 4
US-08-545-809A-110
Sequence 110, Application US/08545809A
Patent No. 6096878
GENERAL INFORMATION:
APPLICANT: Honjo, Tasuku
TITLE OF INVENTION: HUMAN IMMUNOGLOBULIN VH GENE
NUMBER OF INVENTION: SEGMENTS AND DNA FRAGMENTS CONTAINING THE SAME
NUMBER OF SEQUENCES: 145
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson, P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/545,809A
FILING DATE: 27-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP93/00603
FILING DATE: 10-MAY-1993

ATTORNEY/AGENT INFORMATION:
NAME: Freeman, John W.
REGISTRATION NUMBER: 29,066
REFERENCE/DOCKET NUMBER: 06501/004001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-542-5070
TELEFAX: 617-542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 110:
SEQUENCE CHARACTERISTICS:
LENGTH: 117 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-545-809A-110

Query Match
Best Local Similarity 100.0%; Score 31; DB 3; Length 117;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFDP 5
|||||
DB 68 GGFDP 72

RESULT 5
US-08-484-905-110
Sequence 110, Application US/08484905
Patent No. 597651
GENERAL INFORMATION:
APPLICANT: Motiez, Estelle
APPLICANT: Abastado, Jean-Pierre
TITLE OF INVENTION: Kourilsky, Philippe
TITLE OF INVENTION: An Altered Major Histocompatibility
NUMBER OF SEQUENCES: 127
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finegan, Henderson, Farbow, Garrett &
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,905
FILING DATE: 07-JUNE-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/801,818
FILING DATE: 05-DEC-1991
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/792,473
FILING DATE: 15-NOV-1991
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Potter, Jane E. R.
REGISTRATION NUMBER: 33,332
REFERENCE/DOCKET NUMBER: 03495, 0106-03000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4400
TELEFAX: 202-408-4000
INFORMATION FOR SEQ ID NO: 110:
SEQUENCE CHARACTERISTICS:
LENGTH: 255 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: peptide
US-08-484-905-110

Query Match 100.0%; Score 31; DB 2; Length 255;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFPD 5
DB 78 GGFPD 82

RESULT 6
US-08-481-985B-110
Sequence 110, Application US/08481985B
Patent No. 6011146
GENERAL INFORMATION:
APPLICANT: Mottez, Estelle
APPLICANT: Abastado, Jean-Pierre
APPLICANT: Kourilsky, Philippe
TITLE OF INVENTION: Altered Major Histocompatibility Complex
NUMBER OF SEQUENCES: 148
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/481,985B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/801,818
FILING DATE: 05-DEC-1991
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/792,473
FILING DATE: 15-NOV-1991
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 03495,0106-04000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 110:
SEQUENCE CHARACTERISTICS:
LENGTH: 255 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-481-985B-110

Query Match 100.0%; Score 31; DB 3; Length 255;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFPD 5
DB 78 GGFPD 82

RESULT 7
US-08-370-476-110

Sequence 110, Application US/08370476
Patent No. 6153408
GENERAL INFORMATION:
APPLICANT: Mottez, Estelle
APPLICANT: Abastado, Jean-Pierre
APPLICANT: Kourilsky, Philippe
APPLICANT: Lone, Yu-Chun
APPLICANT: Ojcius, David
APPLICANT: Casrouge, Amanda
TITLE OF INVENTION: Altered Major Histocompatibility Complex
NUMBER OF SEQUENCES: 127
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/370,476
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/117,575
FILING DATE: 07-SEP-1993
APPLICATION NUMBER: US 08/072,787
FILING DATE: 06-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/801,818
FILING DATE: 05-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/792,473
FILING DATE: 15-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 05243,0001-01000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 110:
SEQUENCE CHARACTERISTICS:
LENGTH: 255 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-370-476-110

Query Match 100.0%; Score 31; DB 4; Length 255;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFPD 5
DB 78 GGFPD 82

RESULT 8
US-08-047-033-12
Sequence 12, Application US/08047033
Patent No. 5444157
GENERAL INFORMATION:
APPLICANT: Fujio SUZUKI et al.
TITLE OF INVENTION: CHONDROMODULIN-I PROTEIN
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wenderoth, Lind & Ponack

```

STREET: 805 Fifteenth Street, N.W., #700
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 500 KB
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: DisplayWrite
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/047,033
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/821,859
FILING DATE:
APPLICATION NUMBER: US/07/745,497
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX: 202-371-8856
TELEX:
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE:
HYPOHETICAL:
ANTI-SENSE:
FRAGMENT TYPE:
ORIGINAL SOURCE:
ORGANISM:
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL TYPE:
CELL LINE:
ORGANELLE:
IMMEDIATE SOURCE:
LIBRARY:
CLONE:
POSITION IN GENOME:
CHROMOSOME/SEGMENT:
MAP POSITION:
UNITS:
FEATURE:
NAME/KEY:
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS:
TITLE:
JOURNAL:
VOLUME:
ISSUE:
PAGES:
DATE:
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO:
US-08-047-033-12

```

```

Query Match          90.3%; Score 28; DB 1; Length 21;
Best Local Similarity 80.0%; Pred. No. 16;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 GGGDP 5
DB 6 GGYDP 10

```

RESULT 9

```

US-09-413-814-48
Sequence 48, Application US/09413814
Patent No. 6225064
GENERAL INFORMATION:
APPLICANT: Gesellschaft fuer Biotechnologische Forschung mbH
APPLICANT: Bristol-Myers Squibb, Co.
APPLICANT: Beyer, Stefan
APPLICANT: Bloecker, Helmut
APPLICANT: Brandt, Petra
APPLICANT: Cino, Paul M
APPLICANT: Dougherty, Brian A
APPLICANT: Goldberg, Steven L
APPLICANT: Hoffe, Gerhard
APPLICANT: Mueller, Joachim
APPLICANT: Reichenbach, Hans
TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polyketide or
FILE REFERENCE: PCT/US 99/23535
CURRENT APPLICATION NUMBER: US/09/413,814
CURRENT FILING DATE: 1999-10-07
EARLIER APPLICATION NUMBER: DE 198 46 493.2
NUMBER OF SEQ ID NOS: 107
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 48
LENGTH: 2475
TYPE: PRT
ORGANISM: Sorangium cellulosum
US-09-413-814-48

```

```

Query Match          90.3%; Score 28; DB 4; Length 2475;
Best Local Similarity 80.0%; Pred. No. 1.8e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 GGGDP 5
DB 1059 GGYDP 1063

```

RESULT 10

```

US-08-846-134-1
Sequence 1, Application US/08846134
Patent No. 5814481
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Shah, Purni
TITLE OF INVENTION: NOVEL HEAT SHOCK-LIKE PROTEIN
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESS: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/846,134
FILING DATE: Herewith

```

CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Billings, Lucy J.
 REGISTRATION NUMBER: 36,749
 REFERENCE/DOCKET NUMBER: PF-0278 US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-855-0555
 TELEFAX: 415-845-4166
 TELEX:
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 116 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 IMMEDIATE SOURCE:
 LIBRARY: ADRETUT05
 CLONE: 249921
 US-08-846-134-1

Query Match 87.1%; Score 27; DB 2; Length 116;
 Best Local Similarity 80.0%; Pred. No. 1.4e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFPD 5
 DB 52 GGFEF 56

RESULT 11
 5340934-11
 PATENT NO. 5340934
 APPLICANT: TERMINE, JOHN D.; YOUNG, MARIAN F.; FISHER, LARRY W.
 ROBEY, PAMELA G.
 TITLE OF INVENTION: CDNA SEQUENCES OF HUMAN BONE MATRIX PROTEINS
 NUMBER OF SEQUENCES: 13
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/432,044
 FILING DATE: 03-NOV-1989
 SEQ ID NO: 11
 LENGTH: 317
 5340934-11

Query Match 87.1%; Score 27; DB 6; Length 317;
 Best Local Similarity 80.0%; Pred. No. 3.7e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFPD 5
 DB 233 GGFEF 237

RESULT 12
 US-09-470-512A-9
 Sequence 9, Application US/09470512A
 Patent No. 6376652
 GENERAL INFORMATION:
 APPLICANT: PhageTech, Inc.
 TITLE OF INVENTION: Compositions and methods involving an essential Staphylococcus au
 FILE REFERENCE: 21715/1000
 CURRENT APPLICATION NUMBER: US/09/470,512A
 CURRENT FILING DATE: 1999-12-12
 NUMBER OF SEQ ID NOS: 16
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO 9
 LENGTH: 466
 TYPE: PRT
 ORGANISM: Staphylococcus aureus
 US-09-470-512A-9

Query Match 87.1%; Score 27; DB 4; Length 466;
 Best Local Similarity 80.0%; Pred. No. 5.5e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFPD 5
 DB 416 GGFEF 420

RESULT 13
 US-08-887-534A-45
 Sequence 45, Application US/08887534A
 Patent No. 6455323
 GENERAL INFORMATION:
 APPLICANT: Holden, David W.
 TITLE OF INVENTION: ANTI-BACTERIAL METHODS AND MATERIALS
 NUMBER OF SEQUENCES: 106
 CORRESPONDENCE ADDRESS:
 ADDRESSER: Marshall, O'Toole, Gerstein, Murray & Borun
 STREET: 233 South Wacker Drive/6300 Sears Tower
 CITY: Chicago
 STATE: Illinois
 COUNTRY: United States of America
 ZIP: 60606
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/887,534A
 FILING DATE:
 CLASSIFICATION: 536
 ATTORNEY/AGENT INFORMATION:
 NAME: Rin-Laures, Li-Hsien
 REGISTRATION NUMBER: 33,547
 REFERENCE/DOCKET NUMBER: 28341/33996
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (312) 474-6300
 TELEFAX: (312) 474-6448
 INVENTION FOR SEQ ID NO: 45:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 932 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-887-534A-45

Query Match 87.1%; Score 27; DB 4; Length 932;
 Best Local Similarity 80.0%; Pred. No. 1.1e+03;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFPD 5
 DB 798 GGFEF 802

RESULT 14
 US-08-633-770A-1
 Sequence 1, Application US/08633770A
 Patent No. 5908760
 GENERAL INFORMATION:
 APPLICANT: Bojsen, Kirsten
 APPLICANT: Yu, Shukun
 APPLICANT: Kragh, Karsten
 APPLICANT: Christensen, Tove
 APPLICANT: Marcussen, Jan
 TITLE OF INVENTION: ALPHA-1,4-Glucan lyase from a fungus, ITS
 TITLE OF INVENTION: PURIFICATION GENE CLONING AND EXPRESSION IN MICROORGANISMS
 NUMBER OF SEQUENCES: 12
 CORRESPONDENCE ADDRESS:

```

ADDRESS: Knobb, Martens, Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: CA
COUNTRY: U.S.A.
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/633,770A
FILING DATE: July 8, 1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP94/03398
FILING DATE: OCT-15-1994
ATTORNEY/AGENT INFORMATION:
NAME: Altman, Daniel E
REGISTRATION NUMBER: 34,115
REFERENCE/DOCKET NUMBER: DYO06.001APC
TELECOMMUNICATION INFORMATION:
TELEPHONE: 714-760-0404
TELEFAX: 714-760-9502
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1066 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-633-770A-1

Query Match
Best Local Similarity 87.1%; Score 27; DB 2; Length 1066;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFDP 5
Db 666 GGFDP 670

RESULT 15
US-08-633-770A-2
; Sequence 2, Application US/08633770A
; Patent No. 5908760
; GENERAL INFORMATION:
; APPLICANT: Bojsen, Kirsten
; APPLICANT: Yu, Shukun
; APPLICANT: Kragh, Karsten
; APPLICANT: Christensen, Tove
; APPLICANT: Marcussen, Jan
; TITLE OF INVENTION: ALPHA-1,4-GLUCAN LYASE FROM A FUNGUS, ITS
; TITLE OF INVENTION: PURIFICATION GENE CLONING AND EXPRESSION IN MICROORGANISMS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobb, Martens, Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/633,770A
; FILING DATE: July 8, 1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:

```

```

APPLICATION NUMBER: PCT/EP94/03398
FILING DATE: OCT-15-1994
ATTORNEY/AGENT INFORMATION:
NAME: Altman, Daniel E
REGISTRATION NUMBER: 34,115
REFERENCE/DOCKET NUMBER: DYO06.001APC
TELECOMMUNICATION INFORMATION:
TELEPHONE: 714-760-0404
TELEFAX: 714-760-9502
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1070 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-633-770A-2

Query Match
Best Local Similarity 87.1%; Score 27; DB 2; Length 1070;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFDP 5
Db 666 GGFDP 670

Search completed: March 17, 2003, 12:11:50
Job time: 14 secs

```

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: March 17, 2003, 12:10:29 ; Search time 11 seconds
(without alignments)
20.951 Million cell updates/sec

Title: SEQ25-FD
Perfect score: 31
Sequence: 1 gffdp 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 199416 seqs, 46092074 residues
Total number of hits satisfying chosen parameters: 199416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubppa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubppa/PCF_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubppa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubppa/US07_NEW_PUB.pep:*
5: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubppa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubppa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubppa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubppa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31	100.0	56	9 US-09-855-754-9	Sequence 9, Appl1
2	31	100.0	60	9 US-09-866-050A-181	Sequence 181, App
3	31	100.0	60	9 US-09-866-050A-320	Sequence 320, Appl
4	31	100.0	98	9 US-10-194-975-5	Sequence 5, Appl1
5	31	100.0	193	10 US-09-729-674-170	Sequence 170, App
6	31	100.0	271	9 US-09-925-299-978	Sequence 978, App
7	31	100.0	271	10 US-09-925-299-978	Sequence 978, App
8	31	100.0	419	10 US-09-741-669-415	Sequence 415, App
9	31	100.0	419	10 US-09-741-669-415	Sequence 415, App
10	28	90.3	269	10 US-09-795-926-27	Sequence 27, Appl
11	28	90.3	366	10 US-09-795-926-39	Sequence 39, Appl
12	28	90.3	506	10 US-09-795-926-31	Sequence 31, Appl
13	28	90.3	535	10 US-09-795-926-29	Sequence 29, Appl
14	28	90.3	603	10 US-09-795-926-43	Sequence 43, Appl
15	28	90.3	603	12 US-10-001-851-2	Sequence 2, Appl1
16	28	90.3	631	10 US-09-795-926-41	Sequence 41, Appl
17	27	87.1	104	9 US-09-925-299-1155	Sequence 1155, Ap
18	27	87.1	104	10 US-09-925-299-1155	Sequence 1155, Ap
19	27	87.1	618	10 US-09-817-676A-14	Sequence 14, Appl

20	27	87.1	618	10 US-09-870-516-4	Sequence 4, Appl1
21	27	87.1	899	10 US-09-815-242-5356	Sequence 5356, Ap
22	27	87.1	932	10 US-09-815-242-12615	Sequence 12615, A
23	27	87.1	1066	9 US-09-423-126-3	Sequence 3, Appl1
24	27	87.1	1066	10 US-09-280-197-5	Sequence 5, Appl1
25	27	87.1	1070	9 US-09-423-126-4	Sequence 4, Appl1
26	27	87.1	1070	10 US-09-280-197-6	Sequence 6, Appl1
27	26	83.9	232	10 US-09-754-998-2	Sequence 2, Appl1
28	26	83.9	304	9 US-09-738-626-6593	Sequence 6593, Ap
29	26	83.9	354	9 US-09-820-843A-21	Sequence 21, Appl
30	26	83.9	381	9 US-09-893-519A-43	Sequence 43, Appl
31	26	83.9	411	9 US-10-101-664A-982	Sequence 982, App
32	26	83.9	519	9 US-09-738-626-5077	Sequence 5077, Ap
33	25	80.6	11	10 US-09-832-312-51	Sequence 51, Appl
34	25	80.6	20	10 US-09-873-459A-42	Sequence 42, Appl
35	25	80.6	28	10 US-09-739-907-166	Sequence 166, Appl
36	25	80.6	36	10 US-09-864-761-39675	Sequence 39675, A
37	25	80.6	40	10 US-09-864-761-44563	Sequence 44563, A
38	25	80.6	43	12 US-10-042-417-79	Sequence 79, Appl
39	25	80.6	44	10 US-09-864-761-48856	Sequence 48856, A
40	25	80.6	45	10 US-09-864-761-39345	Sequence 39345, A
41	25	80.6	53	9 US-09-796-692-900	Sequence 900, App
42	25	80.6	53	9 US-09-925-299-1537	Sequence 1537, Ap
43	25	80.6	53	9 US-09-925-299-1538	Sequence 1538, Ap
44	25	80.6	53	9 US-09-925-299-1539	Sequence 1539, Ap
45	25	80.6	53	10 US-09-925-302-763	Sequence 763, App

ALIGNMENTS

RESULT 1
US-09-855-754-9
Sequence 9, Application US/09855754
Publication No. US20020192237A1
GENERAL INFORMATION:
APPLICANT: BOURSAX-EBDE, CAROLINE
APPLICANT: GUISO-MACLOPE, NICOLE
TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED
TITLE OF INVENTION: REGIONS OF PERIACETIN IN BORDETELLA PERTUSSIS,
TITLE OF INVENTION: BORDETELLA PARAPERTUSSIS, AND BORDETELLA
TITLE OF INVENTION: BRONCHISEPTICA, THEIR USE IN DIAGNOSTICS, AND IN
FILE REFERENCE: 03495-0206-00000
CURRENT APPLICATION NUMBER: US/09/855,754
CURRENT FILING DATE: 2001-09-10
PRIOR APPLICATION NUMBER: 60/206,969
PRIOR FILING DATE: 2000-05-25
NUMBER OF SEQ ID NOS: 24
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 9
LENGTH: 56
TYPE: PRT
ORGANISM: Bordetella bronchiseptica
US-09-855-754-9
Query Match
Best Local Similarity 100.0%; Score 31; DB 9; Length 56;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGFDP 5
|||||
Db 23 GGFDP 27
RESULT 2
US-09-866-050A-181
Sequence 181, Application US/09866050A
Publication No. US20030040471A1
GENERAL INFORMATION:
APPLICANT: Watson, James D.
APPLICANT: Strachan, Lorna
APPLICANT: Sleeman, Matthew

```
; APPLICANT: Orrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods for Their Use
; FILE REFERENCE: 11000.1011c4u
; CURRENT APPLICATION NUMBER: US/09/866,050A
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 181
; LENGTH: 60
; TYPE: PRT
; ORGANISM: Mouse
US-09-866-050A-181
```

```
Query Match
Best Local Similarity 100.0%; Score 31; DB 9; Length 60;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 GGFD 5
    |||||
Db 51 GGFD 55
```

```
RESULT 3
US-09-866-050A-320
; Sequence 320, Application US/09866050A
; Publication No. US20030040471A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Orrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods for Their Use
; FILE REFERENCE: 11000.1011c4u
; CURRENT APPLICATION NUMBER: US/09/866,050A
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 320
; LENGTH: 60
; TYPE: PRT
; ORGANISM: Mouse
US-09-866-050A-320
```

```
Query Match
Best Local Similarity 100.0%; Score 31; DB 9; Length 60;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 GGFD 5
    |||||
Db 51 GGFD 55
```

```
RESULT 4
US-10-194-975-5
; Sequence 5, Application US/10194975
; Publication No. US20030039649A1
; GENERAL INFORMATION:
; APPLICANT: Foote, Jefferson
; TITLE OF INVENTION: Super Humanized Antibodies
; FILE REFERENCE: 501231.01
; CURRENT APPLICATION NUMBER: US/10/194,975
; CURRENT FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: US 60/305,111
; PRIOR FILING DATE: 2001-07-12
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
```

```
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-194-975-5
```

```
Query Match
Best Local Similarity 100.0%; Score 31; DB 9; Length 98;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 GGFD 5
    |||||
Db 49 GGFD 53
```

```
RESULT 5
US-09-729-674-170
; Sequence 170, Application US/09729674
; Patent No. US20010039335A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: Lavallee, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Agostino, Michael J.
; APPLICANT: Steininger II, Robert J.
; APPLICANT: Spaulding, Vikki
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fechtel, Kim
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYPEPTIDES ENCODING THEM
; FILE REFERENCE: 6055-64X
; CURRENT APPLICATION NUMBER: US/09/729,674
; CURRENT FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: 09/539,330
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 283
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 170
; LENGTH: 193
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-729-674-170
```

```
Query Match
Best Local Similarity 100.0%; Score 31; DB 10; Length 193;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 GGFD 5
    |||||
Db 57 GGFD 61
```

```
RESULT 6
US-09-925-299-978
; Sequence 978, Application US/09925299
; Publication No. US20030040617A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 978
```


LENGTH: 271
TYPE: PRT
ORGANISM: Homo sapiens
US-09-925-299-978

Query Match 100.0%; Score 31; DB 9; Length 271;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFPD 5
Db 60 GGFPD 64

RESULT 7
US-09-925-299-978
Sequence 978, Application US/09925299
Patent No. US20020055627A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA102
CURRENT APPLICATION NUMBER: US/09/925,299
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05883
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
NUMBER OF SEQ ID NOS: 1556
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 978
LENGTH: 271
TYPE: PRT
ORGANISM: Homo sapiens
US-09-925-299-978

Query Match 100.0%; Score 31; DB 10; Length 271;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFPD 5
Db 60 GGFPD 64

RESULT 8
US-09-741-669-415
Sequence 415, Application US/09741669
Patent No. US20020022718A1
GENERAL INFORMATION:
APPLICANT: Forsyth, R. Allyn
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
TITLE OF INVENTION: Genes identified as required for
proliferation of E. coli
FILE REFERENCE: ELITRA.009A
CURRENT APPLICATION NUMBER: US/09/741,669
CURRENT FILING DATE: 2000-12-19
PRIOR APPLICATION NUMBER: US 60/173005
PRIOR FILING DATE: 1999-12-23
NUMBER OF SEQ ID NOS: 481
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 415
LENGTH: 419
TYPE: PRT
ORGANISM: Escherichia coli
US-09-741-669-415

Query Match 100.0%; Score 31; DB 10; Length 419;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFPD 5

Db 386 GGFPD 390

RESULT 9
US-09-741-669-465
Sequence 465, Application US/09741669
Patent No. US20020022718A1
GENERAL INFORMATION:
APPLICANT: Forsyth, R. Allyn
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
TITLE OF INVENTION: Genes identified as required for
proliferation of E. coli
FILE REFERENCE: ELITRA.009A
CURRENT APPLICATION NUMBER: US/09/741,669
CURRENT FILING DATE: 2000-12-19
PRIOR APPLICATION NUMBER: US 60/173005
PRIOR FILING DATE: 1999-12-23
NUMBER OF SEQ ID NOS: 481
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 465
LENGTH: 536
TYPE: PRT
ORGANISM: Escherichia coli
US-09-741-669-465

Query Match 100.0%; Score 31; DB 10; Length 536;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFPD 5
Db 124 GGFPD 128

RESULT 10
US-09-795-926-27
Sequence 27, Application US/09795926
Patent No. US20020098486A1
GENERAL INFORMATION:
APPLICANT: Donoho, Gregory
APPLICANT: Hilbun, Erin
APPLICANT: Turner, C. Alexander Jr.
APPLICANT: Friedrich, Glenn
APPLICANT: Abuin, Alejandro
APPLICANT: Zambrowicz, Brian
APPLICANT: Sands, Arthur T.
APPLICANT: Walke, D. Wade
APPLICANT: Wilganowski, Nathaniel L.
APPLICANT: Hu, Yi
APPLICANT: Kieke, James Allyn
APPLICANT: Potter, David George
TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
POLYNUCLEOTIDES ENCODING THE SAME
FILE REFERENCE: LEX-0144-USA
CURRENT APPLICATION NUMBER: US/09/795,926
CURRENT FILING DATE: 2001-02-28
PRIOR APPLICATION NUMBER: US 60/185,920
PRIOR FILING DATE: 2000-02-29
PRIOR APPLICATION NUMBER: US 60/186,558
PRIOR FILING DATE: 2000-03-02
PRIOR APPLICATION NUMBER: US 60/191,849
PRIOR FILING DATE: 2000-03-24
NUMBER OF SEQ ID NOS: 47
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 27
LENGTH: 269
TYPE: PRT
ORGANISM: homo sapiens
US-09-795-926-27

Query Match 90.3%; Score 28; DB 10; Length 269;

Best Local Similarity 80.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFDP 5
Db 236 GGYDP 240

RESULT 11
US-09-795-926-39

; Sequence 39, Application US/09795926
; Patent No. US20020098486A1

; GENERAL INFORMATION:

; APPLICANT: Donoho, Gregory

; APPLICANT: Hilbun, Erin

; APPLICANT: Turner, C. Alexander Jr.

; APPLICANT: Friederich, Glenn

; APPLICANT: Abulin, Alejandro

; APPLICANT: Zambrowicz, Brian

; APPLICANT: Sands, Arthur T.

; APPLICANT: Walke, D. Wade

; APPLICANT: Wilganowski, Nathaniel L.

; APPLICANT: Hu, Yi

; APPLICANT: Kieke, James Alvin

; APPLICANT: Potter, David George

; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND

; FILE REFERENCE: LEX-0144-USA

; CURRENT APPLICATION NUMBER: US/09/795,926

; PRIOR FILING DATE: 2001-02-28

; PRIOR FILING DATE: 2000-02-29

; PRIOR FILING DATE: 2000-03-02

; PRIOR FILING DATE: 2000-03-02

; PRIOR FILING DATE: 2000-03-24

; NUMBER OF SEQ ID NOS: 47

; SOFTWARE: FastSeq for Windows Version 4.0

; LENGTH: 366

; TYPE: PRT

; ORGANISM: homo sapiens

US-09-795-926-39

Query Match

Best Local Similarity 80.0%; Score 28; DB 10; Length 366;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFDP 5
Db 333 GGYDP 337

RESULT 12
US-09-795-926-31

; Sequence 31, Application US/09795926

; Patent No. US20020098486A1

; GENERAL INFORMATION:

; APPLICANT: Donoho, Gregory

; APPLICANT: Hilbun, Erin

; APPLICANT: Turner, C. Alexander Jr.

; APPLICANT: Friederich, Glenn

; APPLICANT: Abulin, Alejandro

; APPLICANT: Zambrowicz, Brian

; APPLICANT: Sands, Arthur T.

; APPLICANT: Walke, D. Wade

; APPLICANT: Wilganowski, Nathaniel L.

; APPLICANT: Hu, Yi

; APPLICANT: Kieke, James Alvin

; APPLICANT: Potter, David George

; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND

; FILE REFERENCE: LEX-0144-USA

CURRENT APPLICATION NUMBER: US/09/795,926
CURRENT FILING DATE: 2001-02-28

PRIOR APPLICATION NUMBER: US 60/185,920

PRIOR FILING DATE: 2000-02-29

PRIOR APPLICATION NUMBER: US 60/186,558

PRIOR FILING DATE: 2000-03-02

PRIOR APPLICATION NUMBER: US 60/191,849

PRIOR FILING DATE: 2000-03-24

NUMBER OF SEQ ID NOS: 47

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 31

LENGTH: 506

TYPE: PRT

ORGANISM: homo sapiens

US-09-795-926-31

Query Match

Best Local Similarity 80.0%; Score 28; DB 10; Length 506;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFDP 5
Db 236 GGYDP 240

RESULT 13
US-09-795-926-29

; Sequence 29, Application US/09795926

; Patent No. US20020098486A1

; GENERAL INFORMATION:

; APPLICANT: Donoho, Gregory

; APPLICANT: Hilbun, Erin

; APPLICANT: Turner, C. Alexander Jr.

; APPLICANT: Friederich, Glenn

; APPLICANT: Abulin, Alejandro

; APPLICANT: Zambrowicz, Brian

; APPLICANT: Sands, Arthur T.

; APPLICANT: Walke, D. Wade

; APPLICANT: Wilganowski, Nathaniel L.

; APPLICANT: Hu, Yi

; APPLICANT: Kieke, James Alvin

; APPLICANT: Potter, David George

; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND

; FILE REFERENCE: LEX-0144-USA

; CURRENT APPLICATION NUMBER: US/09/795,926

; PRIOR FILING DATE: 2001-02-28

; PRIOR FILING DATE: 2000-02-29

; PRIOR FILING DATE: 2000-02-29

; PRIOR FILING DATE: 2000-03-02

; PRIOR FILING DATE: 2000-03-02

; PRIOR FILING DATE: 2000-03-24

; NUMBER OF SEQ ID NOS: 47

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 29

LENGTH: 535

TYPE: PRT

ORGANISM: homo sapiens

US-09-795-926-29

Query Match

Best Local Similarity 80.0%; Score 28; DB 10; Length 535;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFDP 5
Db 236 GGYDP 240

RESULT 14
US-09-795-926-43

; Sequence 43, Application US/09795926

Search completed: March 17, 2003, 12:21:33
Job time : 12 secs

```

Patent No. US20020098486A1
GENERAL INFORMATION:
APPLICANT: Donoho, Gregory
APPLICANT: Hilbun, Erin
APPLICANT: Turner, C. Alexander Jr.
APPLICANT: Friedrich, Glenn
APPLICANT: Abuin, Alejandro
APPLICANT: Zambrowicz, Brian
APPLICANT: Sands, Arthur T.
APPLICANT: Walke, D. Wade
APPLICANT: Wilgenowski, Nathaniel L.
APPLICANT: Hu, Yi
APPLICANT: Kieke, James Alvin
APPLICANT: Potter, David George
TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
FILE REFERENCE: LEX-0144-USA
CURRENT APPLICATION NUMBER: US/09/795,926
CURRENT FILING DATE: 2001-02-28
PRIOR APPLICATION NUMBER: US 60/185,920
PRIOR FILING DATE: 2000-02-29
PRIOR APPLICATION NUMBER: US 60/186,558
PRIOR FILING DATE: 2000-03-02
PRIOR APPLICATION NUMBER: US 60/191,849
PRIOR FILING DATE: 2000-03-24
NUMBER OF SEQ ID NOS: 47
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 43
LENGTH: 603
TYPE: PRT
ORGANISM: homo sapiens
US-09-795-926-43
    
```

```

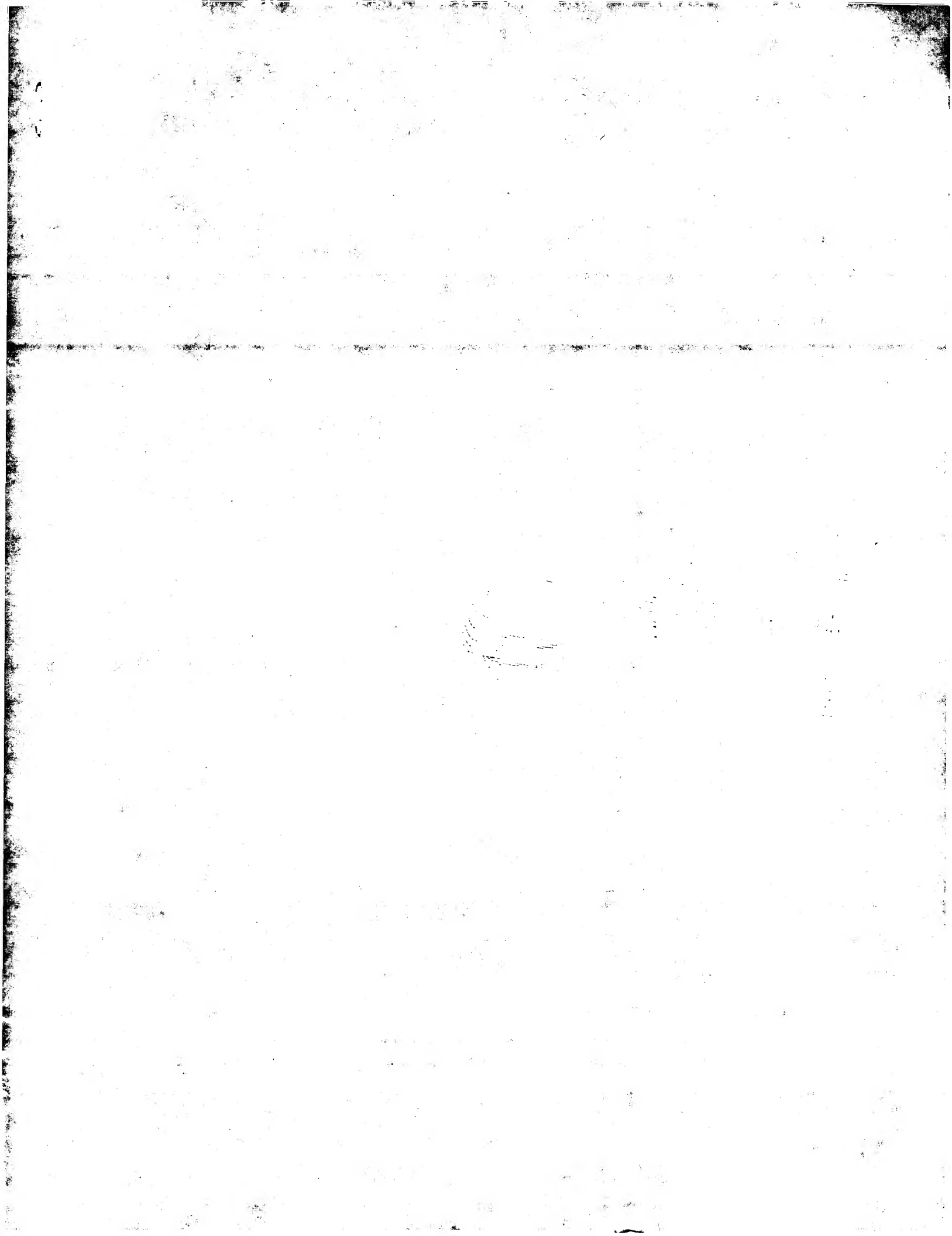
Query Match          90.3%; Score 28; DB 10; Length 603;
Best Local Similarity 80.0%; Pred. No. 4e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 1 GGEDP 5
Db 333 GGYDP 337
    
```

```

RESULT 15
US-10-001-851-2
Sequence 2, Application US/10001851
Patent No. US20020115628A1
GENERAL INFORMATION:
APPLICANT: MEYERS, Rachel A.
APPLICANT: WILKINSON, Mark
TITLE OF INVENTION: 47169 and 33935, No. US20020115628A1 Human Glycosyl Transferase
FILE REFERENCE: 10147-56U1
CURRENT APPLICATION NUMBER: US/10/001,851
CURRENT FILING DATE: 2001-11-20
PRIOR APPLICATION NUMBER: US 60/249,939
PRIOR FILING DATE: 2000-11-20
NUMBER OF SEQ ID NOS: 29
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 603
TYPE: PRT
ORGANISM: Homo sapiens
US-10-001-851-2
    
```

```

Query Match          90.3%; Score 28; DB 12; Length 603;
Best Local Similarity 80.0%; Pred. No. 4e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 1 GGEDP 5
Db 333 GGYDP 337
    
```



GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 17, 2003, 12:06:28 ; Search time 12.3333 Seconds
(without alignments)
38.973 Million cell updates/sec

Title: SEQ25-FD
Perfect score: 31
Sequence: 1 gfdp 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31	100.0	62	2	136967 gene MHC DO-alpha
2	31	100.0	62	2	137040 gene MHC DO-alpha
3	31	100.0	62	2	161791 gene MHC DO-alpha
4	31	100.0	62	2	161808 gene COX4P1 proteol
5	31	100.0	62	2	161804 gene MHC DO-alpha
6	31	100.0	62	2	161797 gene MHC DO-alpha
7	31	100.0	62	2	161795 gene MHC DO-alpha
8	31	100.0	62	2	161800 gene MHC DO-alpha
9	31	100.0	62	2	161790 gene MHC DO-alpha
10	31	100.0	62	2	161794 gene MHC DO-alpha
11	31	100.0	63	2	136900 gene MHC DO-alpha
12	31	100.0	63	2	161787 gene MHC DO-alpha
13	31	100.0	63	2	161801 gene MHC DO-alpha
14	31	100.0	63	2	161789 gene MHC DO-alpha
15	31	100.0	63	2	161798 gene MHC DO-alpha
16	31	100.0	63	2	161798 gene MHC DO-alpha
17	31	100.0	63	2	161798 gene MHC DO-alpha
18	31	100.0	63	2	161798 gene MHC DO-alpha
19	31	100.0	63	2	161798 gene MHC DO-alpha
20	31	100.0	63	2	161798 gene MHC DO-alpha
21	31	100.0	63	2	161798 gene MHC DO-alpha
22	31	100.0	63	2	161798 gene MHC DO-alpha
23	31	100.0	63	2	161798 gene MHC DO-alpha
24	31	100.0	63	2	161798 gene MHC DO-alpha
25	31	100.0	63	2	161798 gene MHC DO-alpha
26	31	100.0	63	2	161798 gene MHC DO-alpha
27	31	100.0	63	2	161798 gene MHC DO-alpha
28	31	100.0	63	2	161798 gene MHC DO-alpha
29	31	100.0	63	2	161798 gene MHC DO-alpha

ALIGNMENTS

30 31 100.0 127 2 PH1414 Ig heavy chain V r
31 31 100.0 127 2 PH1415 Ig heavy chain V r
32 31 100.0 138 2 PN0538 Ig heavy chain V r
33 31 100.0 151 2 PL0011 Ig heavy chain pre
34 31 100.0 166 2 JC4664 T-cell receptor ze
35 31 100.0 166 2 JC4664 T-cell surface gly
36 31 100.0 168 2 S64830 hypothetical prote
37 31 100.0 210 2 A71876 oxygen-insensitive
38 31 100.0 216 2 B64639 fibron protein proteu
39 31 100.0 242 2 A36148 cell surface glyco
40 31 100.0 243 2 A54290 flagellar L-ring p
41 31 100.0 244 2 A87505 MHC class II histo
42 31 100.0 255 1 HLH0D1 tetraspan TSPAN-5
43 31 100.0 264 2 A59261 hypothetical prote
44 31 100.0 264 2 E91158
45 31 100.0 264 2 E91158

RESULT 1
136967
gene MHC DO-alpha 1 protein - chimpanzee (fragment)
C:Species: Pan troglodytes (chimpanzee)
C>Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 23-Jul-1999
C:Accession: 136967
R:Kenter, M.; Otting, N.; Anholts, J.; Leunissen, J.A.M.; Jonker, M.; Bontrop, R.E.
Immunogenetics 36, 71-78, 1992
A:Title: Evolutionary relationships among the primate MHC-DQA1 and DQA2 alleles.
A:Reference number: 136899; PMID:92307745; PMID:1612647
A:Accession: 136967
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-62 <RES>
A:Cross-references: GB:M76189; NID:q176840; PIDN:AAA35428.1; PID:9553156
C:Superfamily: class II histocompatibility antigen; immunoglobulin homology

Query Match 100.0%; Score 31; DB 2; Length 62;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFDP 5
DB 38 GGFDP 42

RESULT 2
137040
gene MHC DO-alpha 1 protein - common gibbon (fragment)
C:Species: Hylobates lar (common gibbon, white-handed gibbon)
C>Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 23-Jul-1999
R:Kenter, M.; Otting, N.; Anholts, J.; Leunissen, J.A.M.; Jonker, M.; Bontrop, R.E.
Immunogenetics 36, 71-78, 1992
A:Title: Evolutionary relationships among the primate MHC-DQA1 and DQA2 alleles.
A:Reference number: 136899; PMID:92307745; PMID:1612647
A:Accession: 137040
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-62 <RES>
A:Cross-references: GB:M76212; NID:q177024; PIDN:AAA35459.1; PID:9553157
C:Superfamily: class II histocompatibility antigen; immunoglobulin homology

Query Match 100.0%; Score 31; DB 2; Length 62;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFDP 5
DB 38 GGFDP 42

RESULT 3

161791

gene MHC DQ-alpha 1 protein - rhesus macaque (fragment)

C:Species: Macaca mulatta (rhesus macaque)

C:Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 23-Jul-1999

C:Accession: 161791

R:Kenter, M.; Otting, N.; Anholts, J.; Leunissen, J.A.M.; Jonker, M.; Bontrop, R.E.

Immunogenetics 36, 71-78, 1992

A:Title: Evolutionary relationships among the primate MHC-DQA1 and DQA2 alleles.

A:Reference number: 136899; MUID:92307745; PMID:1612647

A:Accession: 161791

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-62 <RES>

A:Cross-references: GB:M76201; NID:9342139; PIDN:AAA36859.1; PID:9553823

C:Superfamily: class II histocompatibility antigen; immunoglobulin homology

Query Match

Best Local Similarity 100.0%; Score 31; DB 2; Length 62;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFDP 5

DB 38 GGFDP 42

RESULT 4

161808

gene COX4P1 protein - orangutan (fragment)

C:Species: Pongo pygmaeus (orangutan)

C:Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 23-Jul-1999

C:Accession: 161808

R:Kenter, M.; Otting, N.; Anholts, J.; Leunissen, J.A.M.; Jonker, M.; Bontrop, R.E.

Immunogenetics 36, 71-78, 1992

A:Title: Evolutionary relationships among the primate MHC-DQA1 and DQA2 alleles.

A:Reference number: 136899; MUID:92307745; PMID:1612647

A:Accession: 161808

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-62 <RES>

A:Cross-references: GB:M76223; NID:9342859; PIDN:AAA36940.1; PID:9553836

C:Superfamily: class II histocompatibility antigen; immunoglobulin homology

Query Match

Best Local Similarity 100.0%; Score 31; DB 2; Length 62;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFDP 5

DB 38 GGFDP 42

RESULT 5

161804

gene MHC DO-alpha 1 protein - orangutan (fragment)

C:Species: Pongo pygmaeus (orangutan)

C:Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 23-Jul-1999

C:Accession: 161804

R:Kenter, M.; Otting, N.; Anholts, J.; Leunissen, J.A.M.; Jonker, M.; Bontrop, R.E.

Immunogenetics 36, 71-78, 1992

A:Title: Evolutionary relationships among the primate MHC-DQA1 and DQA2 alleles.

A:Reference number: 136899; MUID:92307745; PMID:1612647

A:Accession: 161804

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-62 <RES>

A:Cross-references: GB:M76218; NID:9342851; PIDN:AAA36936.1; PID:9553834

C:Superfamily: class II histocompatibility antigen; immunoglobulin homology

Query Match

Best Local Similarity 100.0%; Score 31; DB 2; Length 62;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFDP 5

DB 38 GGFDP 42

RESULT 6

161797

gene MHC DQ-alpha 1 protein - crab-eating macaque (fragment)

C:Species: Macaca fascicularis (crab-eating macaque)

C:Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 23-Jul-1999

C:Accession: 161797

R:Kenter, M.; Otting, N.; Anholts, J.; Leunissen, J.A.M.; Jonker, M.; Bontrop, R.E.

Immunogenetics 36, 71-78, 1992

A:Title: Evolutionary relationships among the primate MHC-DQA1 and DQA2 alleles.

A:Reference number: 136899; MUID:92307745; PMID:1612647

A:Accession: 161797

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-62 <RES>

A:Cross-references: GB:M76207; NID:9342151; PIDN:AAA36865.1; PID:9553826

C:Superfamily: class II histocompatibility antigen; immunoglobulin homology

Query Match

Best Local Similarity 100.0%; Score 31; DB 2; Length 62;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFDP 5

DB 38 GGFDP 42

RESULT 7

161795

gene MHC DO-alpha 1 protein - stump-tailed macaque (fragment)

C:Species: Macaca arctoides (stump-tailed macaque)

C:Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 23-Jul-1999

C:Accession: 161795

R:Kenter, M.; Otting, N.; Anholts, J.; Leunissen, J.A.M.; Jonker, M.; Bontrop, R.E.

Immunogenetics 36, 71-78, 1992

A:Title: Evolutionary relationships among the primate MHC-DQA1 and DQA2 alleles.

A:Reference number: 136899; MUID:92307745; PMID:1612647

A:Accession: 161795

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-62 <RES>

A:Cross-references: GB:M76205; NID:9342147; PIDN:AAA36863.1; PID:9553825

C:Superfamily: class II histocompatibility antigen; immunoglobulin homology

Query Match

Best Local Similarity 100.0%; Score 31; DB 2; Length 62;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFDP 5

DB 38 GGFDP 42

RESULT 8

161800

gene MHC DO-alpha 1 protein - rhesus macaque (fragment)

C:Species: Macaca mulatta (rhesus macaque)

C:Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 23-Jul-1999

C:Accession: 161800

R:Kenter, M.; Otting, N.; Anholts, J.; Leunissen, J.A.M.; Jonker, M.; Bontrop, R.E.

Immunogenetics 36, 71-78, 1992

A:Title: Evolutionary relationships among the primate MHC-DQA1 and DQA2 alleles.

A:Reference number: 136899; MUID:92307745; PMID:1612647

A:Accession: 161800

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-62 <RES>

A:Cross-references: GB:M76227; NID:9342157; PIDN:AAA36868.1; PID:9553827

C:Superfamily: class II histocompatibility antigen; immunoglobulin homology

Query Match 100.0%; Score 31; DB 2; Length 62;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFDP 5
 |||||
 Db 38 GGFDP 42

RESULT 9

161790

gene MHC DO-alpha 1 protein - rhesus macaque (fragment)
 C:Species: Macaca mulatta (rhesus macaque)
 C>Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 23-Jul-1999
 C:Accession: 161790

R:Kenter, M.; Otting, N.; Anholts, J.; Leunissen, J.A.M.; Jonker, M.; Bontrop, R.E.
 Immunogenetics 36, 71-78, 1992
 A:Title: Evolutionary relationships among the primate MHC-DQA1 and DQA2 alleles.
 A:Reference number: I36899; MUID:92307745; PMID:1612647
 A:Accession: 161790

A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-62 <RES>
 A:Cross-references: GB:M76200; NID:g342137; PIDN:AAA36858.1; PID:g553822
 C:Superfamily: class II histocompatibility antigen; immunoglobulin homology

Query Match 100.0%; Score 31; DB 2; Length 62;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFDP 5
 |||||
 Db 38 GGFDP 42

RESULT 10

161794

gene MHC DO-alpha 1 protein - stump-tailed macaque (fragment)
 C:Species: Macaca arctoides (stump-tailed macaque)
 C>Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 23-Jul-1999
 C:Accession: 161794

R:Kenter, M.; Otting, N.; Anholts, J.; Leunissen, J.A.M.; Jonker, M.; Bontrop, R.E.
 Immunogenetics 36, 71-78, 1992
 A:Title: Evolutionary relationships among the primate MHC-DQA1 and DQA2 alleles.
 A:Reference number: I36899; MUID:92307745; PMID:1612647
 A:Accession: 161794

A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-62 <RES>
 A:Cross-references: GB:M76204; NID:g342145; PIDN:AAA36862.1; PID:g553824
 C:Superfamily: class II histocompatibility antigen; immunoglobulin homology

Query Match 100.0%; Score 31; DB 2; Length 62;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFDP 5
 |||||
 Db 38 GGFDP 42

RESULT 11

136900

gene MHC DO-alpha 1 protein - green monkey (fragment)
 C:Species: Cercopithecus aethiops (green monkey, grivet)
 C>Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 23-Jul-1999
 C:Accession: 136900

R:Kenter, M.; Otting, N.; Anholts, J.; Leunissen, J.A.M.; Jonker, M.; Bontrop, R.E.
 Immunogenetics 36, 71-78, 1992
 A:Title: Evolutionary relationships among the primate MHC-DQA1 and DQA2 alleles.
 A:Reference number: I36899; MUID:92307745; PMID:1612647
 A:Accession: 136900

A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-63 <RES>
 A:Cross-references: GB:M76187; NID:g176517; PIDN:AAA35364.1; PID:g176518
 C:Superfamily: class II histocompatibility antigen; immunoglobulin homology

A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-63 <RES>
 A:Cross-references: GB:M76187; NID:g176517; PIDN:AAA35364.1; PID:g176518
 C:Superfamily: class II histocompatibility antigen; immunoglobulin homology

Query Match 100.0%; Score 31; DB 2; Length 63;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFDP 5
 |||||
 Db 38 GGFDP 42

RESULT 12

161787

gene MHC DO-alpha 1 protein - rhesus macaque (fragment)
 C:Species: Macaca mulatta (rhesus macaque)
 C>Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 23-Jul-1999
 C:Accession: 161787

R:Kenter, M.; Otting, N.; Anholts, J.; Leunissen, J.A.M.; Jonker, M.; Bontrop, R.E.
 Immunogenetics 36, 71-78, 1992
 A:Title: Evolutionary relationships among the primate MHC-DQA1 and DQA2 alleles.
 A:Reference number: I36899; MUID:92307745; PMID:1612647
 A:Accession: 161787

A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-63 <RES>
 A:Cross-references: GB:M76195; NID:g342131; PIDN:AAA36855.1; PID:g342132
 C:Superfamily: class II histocompatibility antigen; immunoglobulin homology

Query Match 100.0%; Score 31; DB 2; Length 63;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFDP 5
 |||||
 Db 38 GGFDP 42

RESULT 13

161801

gene MHC DO-alpha 1 protein - rhesus macaque (fragment)
 C:Species: Macaca mulatta (rhesus macaque)
 C>Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 23-Jul-1999
 C:Accession: 161801

R:Kenter, M.; Otting, N.; Anholts, J.; Leunissen, J.A.M.; Jonker, M.; Bontrop, R.E.
 Immunogenetics 36, 71-78, 1992
 A:Title: Evolutionary relationships among the primate MHC-DQA1 and DQA2 alleles.
 A:Reference number: I36899; MUID:92307745; PMID:1612647
 A:Accession: 161801

A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-63 <RES>
 A:Cross-references: GB:M76228; NID:g342159; PIDN:AAA36869.1; PID:g442160
 C:Superfamily: class II histocompatibility antigen; immunoglobulin homology

Query Match 100.0%; Score 31; DB 2; Length 63;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFDP 5
 |||||
 Db 38 GGFDP 42

RESULT 14

161789

gene MHC DO-alpha 1 protein - rhesus macaque (fragment)
 C:Species: Macaca mulatta (rhesus macaque)
 C>Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 23-Jul-1999
 C:Accession: 161789

A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-63 <RES>
 A:Cross-references: GB:M76187; NID:g176517; PIDN:AAA35364.1; PID:g176518
 C:Superfamily: class II histocompatibility antigen; immunoglobulin homology

R:Kenter, M.; Otting, N.; Anholts, J.; Leunissen, J.A.M.; Jonker, M.; Bontrop, R.E.
 Immunogenetics 36, 71-76, 1992
 A:Title: Evolutionary relationships among the primate MHC-DQA1 and DQA2 alleles.
 A:Reference number: I36899; MUID:92307745; PMID:1612647
 A:Accession: I61789
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-63 <RES>
 A:Cross-references: GB:M76199; NID:g342135; PIDN:AAA36857.1; PID:g342136
 C:Superfamily: class II histocompatibility antigen; immunoglobulin homology

Query Match 100.0%; Score 31; DB 2; Length 63;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFPD 5
 |||||
 DB 38 GGFPD 42

RESULT 15

I61798
 gene MHC DQ-alpha 1 protein - crab-eating macaque (fragment)
 C:Species: Macaca fascicularis (crab-eating macaque)
 C:Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 23-Jul-1999
 C:Accession: I61798
 R:Kenter, M.; Otting, N.; Anholts, J.; Leunissen, J.A.M.; Jonker, M.; Bontrop, R.E.
 Immunogenetics 36, 71-76, 1992
 A:Title: Evolutionary relationships among the primate MHC-DQA1 and DQA2 alleles.
 A:Reference number: I36899; MUID:92307745; PMID:1612647
 A:Accession: I61798
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-63 <RES>
 A:Cross-references: GB:M76208; NID:g342153; PIDN:AAA36866.1; PID:g342154
 C:Superfamily: class II histocompatibility antigen; immunoglobulin homology

Query Match 100.0%; Score 31; DB 2; Length 63;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFPD 5
 |||||
 DB 38 GGFPD 42

Search completed: March 17, 2003, 12:11:05
 Job time : 13.3333 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 17, 2003, 12:05:19 ; Search time 7 seconds
(Without alignments)

29.626 Million cell updates/sec

Title: SEQ25-FD
Perfect score: 31
Sequence: 1 gffdp 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	31	100.0	Y005_VIBCH	Q9KvY3 vibrio chol
2	31	100.0	Y1DD_HAEIN	P4972 haemophilus
3	31	100.0	Y150_DEIRA	Q9RS4 delinococcus
4	31	100.0	CD32_SHEEP	P29329 ovis aries
5	31	100.0	FLGH_CACCR	P22606 caulobacter
6	31	100.0	HA25_HUMAN	P01907 homo sapien
7	31	100.0	YRPD_ECOLI	P45733 escherichia
8	31	100.0	PSBP_CUCSA	Q9S148 cucumis sat
9	31	100.0	PSBP_FRIG	Q9S080 fitillaria
10	31	100.0	T4S9_HUMAN	O60628 homo sapien
11	31	100.0	F2FL_YEAST	P32805 saccharomyc
12	31	100.0	MCH_XANAU	O8R01 xanthobacte
13	31	100.0	DICH_DROME	Q24533 drosophila
14	31	100.0	ASSY_SCHPO	Q94334 schizosacch
15	31	100.0	GLPB_ECOLI	P13033 escherichia
16	31	100.0	DCMD_METTE	O50538 methanosarc
17	31	100.0	EF11_XENLA	P17506 xenopus lae
18	31	100.0	THSB_SULAC	O9V214 sulfolobus
19	31	100.0	YAGH_ECOLI	P77713 escherichia
20	31	100.0	VPP_BPHPI	P51718 bacterioph
21	31	100.0	HS70_BLAEM	P48720 blastoclad
22	31	100.0	CHST_USTMA	Q9J127 ustiliag m
23	31	100.0	SOXA_RHIME	O87386 rhizobium
24	31	100.0	NIT4_NEUCR	P28349 neurospora
25	28	90.3	Y331_SYNY3	O55785 synecocyst
26	28	90.3	YE62_THEMA	O9X133 thermotoga
27	28	90.3	YIDD_PSEPU	P25753 pseudomonas
28	28	90.3	YB64_PASMU	Q9C143 pasteurella
29	28	90.3	YK45_PSEAE	O91270 pseudomonas
30	28	90.3	Y989_XYLEAE	O9P29 xylella fas
31	28	90.3	AMIB_BUCAI	P57638 buchiera ap
32	28	90.3	YI23_MYCTU	O50610 mycobacteri
33	28	90.3	YB94_MERTH	O27262 methanobact

ALIGNMENTS

RESULT 1			
ID	Y005_VIBCH	STANDARD:	PRT: 85 AA.
AC	Q9KvY3:		
DT	16-OCT-2001 (Rel. 40, Created)		
DT	16-OCT-2001 (Rel. 40, Last sequence update)		
DT	16-OCT-2001 (Rel. 40, Last annotation update)		
DE	Hypothetical protein VC0005.		
GN	VC0005.		
OS	Vibrio cholerae.		
OC	Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.		
OX	NCBI_TaxID=666;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=El Tor N16961 / Serotype O1;		
RX	MEDLINE=20406833; PubMed=10952301;		
RA	Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,		
RA	Dodson R.J., Hail D.H., Hickey E.K., Peterson J.D., Umayam L.A.,		
RA	Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,		
RA	Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragol I., Sellers P.,		
RA	McDonald L., Uitterback T., Fleischmann R.D., Nierman W.C., White O.,		
RA	Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,		
RA	Frazer C.M.;		
RT	"DNA sequence of both chromosomes of the cholera pathogen Vibrio		
RT	cholerae.";		
RL	Nature 406:477-483(2000).		
CC	-I- SIMILARITY: BELONGS TO THE UPF0161 FAMILY.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/		
CC	or send an email to license@isb-sib.ch).		
CC	-----		
DR	EMBL: AE004093; AAF93183.1; -		
DR	TIGR: VC0005; -		
DR	InterPro: IPR002696; DUF37.		
DR	Pfam: PF01809; DUF37.1.		
DR	Prodom: PD004225; DUF37.1.		
DR	TIGRFAMS: TIGR00278; DUF37.1.		
DR	Hypothetical protein; Complete proteome.		
DR	SEQUENCE 85 AA; 9616 MW; 0289CE6770E315A9 CRC64;		
SQ			
Query Match			
Best Local Similarity 100.0%; Score 31; DB 1; Length 85;			
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
OY 1 GGFDP 5			
Db 70 GGFDP 74			
RESULT 2			
Y1DD_HAEIN			

```

ID Y1DD_HAEIN STANDARD: PRT: 86 AA.
AC P44972;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein H11000.
GN H11000.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Rd / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kierulff A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrman J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RD Rd."
RL Science 269:496-512(1995).
CC -1 SIMILARITY: BELONGS TO THE UPF0161 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC -----
CC EMBL: U32781; AAC22662.1; -
CC TIGR: H11000; -
CC DR InterPro: IPR002696; DUF37.
CC DR Pfam: PF01809; DUF37.1.
CC DR ProDom: PD004225; DUF37.1.
CC DR TIGRPFAMs: TIGR00278; DUF37.1.
CC KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 86 AA; 9564 MW; 7AD4FA5A27020AB CRC64;

Query Match
Best Local Similarity 100.0%; Score 31; DB 1; Length 86;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFDP 5
DB 70 GGFDP 74

RESULT 3
Y150_DEIRA STANDARD: PRT: 96 AA.
AC Q9RSH4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein DR2150.
GN DR2150.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-R1;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,

```

```

RA Dodson R.J., Haft D.H., Gwin M.L., Nelson W.C., Richardson D.L.,
RA Morfitt K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Uterback T., Zalewski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
RD radiodurans R1."
RL Science 286:1571-1577(1999).
CC -1 SIMILARITY: BELONGS TO THE UPF0161 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC -----
CC EMBL: AF002049; AF11698.1; -
CC TIGR: DR2150; -
CC DR InterPro: IPR002696; DUF37.
CC DR Pfam: PF01809; DUF37.1.
CC DR ProDom: PD004225; DUF37.1.
CC DR TIGRPFAMs: TIGR00278; DUF37.1.
CC KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 96 AA; 10628 MW; EA1201670EE9D1E3 CRC64;

Query Match
Best Local Similarity 100.0%; Score 31; DB 1; Length 96;
Matches 5; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

OY 1 GGFDP 5
DB 65 GGFDP 69

RESULT 4
CD3Z_SHEEP STANDARD: PRT: 166 AA.
AC P29329;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE T-cell surface glycoprotein CD3 zeta chain precursor (T-cell receptor
DE T3 zeta chain).
GN CD3Z.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=White alpine;
RX MEDLINE=9313105; PubMed=8420837;
RA Hehn W.R., Tunncliffe A.;
RT "Invariant components of the sheep T-cell antigen receptor: cloning
RT of the CD3 epsilon and Tcr zeta chains."
RL Immunogenetics 37:279-284(1993).
CC -1 FUNCTION: PROBABLE ROLE IN ASSEMBLY AND EXPRESSION OF THE TCR
CC COMPLEX AS WELL AS SIGNAL TRANSDUCTION UPON ANTIGEN TRIGGERING.
CC -1 SUBUNIT: THE TCR/CD3 COMPLEX OF T LYMPHOCYTES CONSISTS OF EITHER A
CC TCR ALPHA/BETA OR TCR GAMMA/DELTA HETERODIMER COEXPRESSED AT THE
CC CELL SURFACE WITH THE INVARIANT SUBUNITS OF CD3 LABELED GAMMA,
CC DELTA, EPSILON, ZETA, AND ETA.
CC -1 SUBCELLULAR LOCATION: Type I membrane protein.
CC -1 PMT: PHOSPHORYLATED ON TIR RESIDUES AFTER T-CELL RECEPTOR
CC TRIGGERING (BY SIMILARITY).
CC -1 SIMILARITY: BELONGS TO THE CD3Z/FCER1G FAMILY.
CC -1 SIMILARITY: CONTAINS 3 ITAM DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Z12968; CAA78312.1; -
 DR PIR: S22980; S22980.
 DR InterPro: IPR003110; ITAM.
 DR Pfam: PF02189; ITAM; 3.
 DR SMART: SM00077; ITAM; 3.
 KM T-cell: Receptor; Transmembrane; Signal; Repeat; Phosphorylation.
 FT SIGNAL 1 21 BY SIMILARITY
 FT CHAIN 22 166 T-CELL SURFACE GLYCOPROTEIN CD3 ZETA
 FT CHAIN
 FT DOMAIN 22 30 POTENTIAL.
 FT TRANSMEM 31 51 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 52 166 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 69 89 ITAM 1.
 FT DOMAIN 108 129 ITAM 2.
 FT DOMAIN 141 161 ITAM 3.
 FT DISULFID 32 32 INTERCHAIN (POTENTIAL).
 FT MOD_RES 155 155 PHOSPHORYLATION (MAY HAVE A REGULATORY
 FT ROLE) (BY SIMILARITY).
 SQ SEQUENCE 166 AA; 18704 MW; E7D89AD84E5831A CRC64;
 Query Match 100.0%; Score 31; DB 1; Length 166;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 GGDP 5
 Db 90 GGDP 94
 RESULT 5
 FLGH CAUCR STANDARD; PRT; 244 AA.
 ID FLGH CAUCR
 AC P22606;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Flagellar L-ring protein precursor (Basal body L-ring protein).
 GN FLGH OR FLBN OR CC2066.
 OS Caulobacter crescentus.
 OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
 OC Caulobacter.
 OX NCBI_TaxID=155892;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 19089 / CB15;
 RC MEDLINE-91008987; Pubmed-2211524;
 RA Dingwall A., Guber J.W., Shapiro L.,
 RT "Identification of a Caulobacter basal body structural gene and a
 RT cis-acting site required for activation of transcription.";
 RL J. Bacteriol. 172:6066-6076(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 19089 / CB15;
 RC MEDLINE-21173688; Pubmed-11259647;
 RA Nieman W.C., Feldblum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
 RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
 RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
 RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwin M.L., Haft D.H.,
 RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
 RA Ueberall J., Tran K., Wolf A., Yamathavan J., Esmolaeva M., White O.,
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
 RT Complete genome sequence of Caulobacter crescentus.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
 CC -1- FUNCTION: ASSEMBLES AROUND THE ROD TO FORM THE L-RING AND PROBABLY
 CC PROTECTS THE MOTOR/BASAL BODY FROM SHEARING FORCES DURING
 CC ROTATION.

CC -1- SUBUNIT: THE BASAL BODY CONSTITUTES A MAJOR PORTION OF THE
 CC FLAGELLAR ORGANELLE AND CONSISTS OF FIVE RINGS (E,L,P,S, AND M)
 CC MOUNTED ON A CENTRAL ROD.
 CC -1- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid
 CC anchor (Probable).
 CC -1- SIMILARITY: BELONGS TO THE FLGH FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M37270; AAA23035.1; -
 DR PIR: A6005880; AAK24037.1; -
 DR PIR: A36148; A36148.
 DR TIGR: CC2066; -
 DR InterPro: IPR000527; Flag_Lring.
 DR InterPro: IPR000437; Prok_Lipoprot.
 DR Pfam: PF02107; FLGH; 1.
 DR PRINTS: PRO1008; FLAGRINGFLGH.
 DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; FALSE NEG.
 KM Flagella; Outer membrane; Lipoprotein; Signal; Complete proteome.
 FT SIGNAL 1 18 POTENTIAL.
 FT CHAIN 19 244 FLAGELLAR L-RING PROTEIN.
 FT LIPID 19 19 N-ACYL DIGLYCERIDE (PROBABLE).
 FT CONFLICT 57 69 AASANSIMRVGAR -> RFGOLAVAGRGP (IN REF.
 FT 1).
 FT CONFLICT 95 95 S -> R (IN REF. 1).
 FT CONFLICT 98 99 NA -> KP (IN REF. 1).
 FT CONFLICT 113 113 P -> R (IN REF. 1).
 FT CONFLICT 150 150 MISSING (IN REF. 1).
 FT CONFLICT 167 167 V -> L (IN REF. 1).
 SQ SEQUENCE 244 AA; 25525 MW; 7C08B8ACF105BA30 CRC64;
 Query Match 100.0%; Score 31; DB 1; Length 244;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 GGDP 5
 Db 128 GGDP 132
 RESULT 6
 HA25_HUMAN STANDARD; PRT; 255 AA.
 ID HA25_HUMAN
 AC P01907;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE HLA class II histocompatibility antigen, DQ(5) alpha chain precursor
 DE (DC-1 alpha chain).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (CLONE HB20).
 RX MEDLINE-84168117; Pubmed-6584734;
 RA Affray C., Lillie J.W., Arnot D., Grossberger D., Kappes D.,
 RA Strominger J.L.;
 RT "Isotypic and allotypic variation of human class II
 RT histocompatibility antigen alpha-chain genes.";
 RL Nature 308:327-333(1984).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 DR EMBL; X00452; CA25141.1; -
 DR PIR; A02211; HLHDL.
 DR HSSP; P01910; ITRK.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR003597; Ig_c1.
 DR InterPro; IPR001003; MHC_II_alpha.
 DR Pfam; PF00047; Ig; 1.
 DR Pfam; PF00993; MHC_II_alpha; 1.
 DR SMART; SM00407; IgC1; 1.
 DR PROSITE; PS00290; IG_MHC; 1.
 KM MHC II; Transmembrane; Glycoprotein; Signal.
 FT SIGNAL 1 23
 FT CHAIN 24 255
 FT 255
 FT DOMAIN 24 110
 FT DOMAIN 111 204
 FT DOMAIN 205 217
 FT TRANSMEM 218 240
 FT DOMAIN 241 255
 FT DISULFID 133 189
 FT CARBOHYD 104 104
 FT CARBOHYD 144 144
 FT SEQUENCE 255 AA; 28058 MW; 2E33F13105919571 CRC64;
 N-LINKED (GICNAC...) (POTENTIAL).
 N-LINKED (GICNAC...) (POTENTIAL).
 CYTOPLASMIC TAIL.
 BY SIMILARITY.
 DO(S) ALPHA CHAIN.
 EXTRACELLULAR ALPHA-1.
 EXTRACELLULAR ALPHA-2.
 CONNECTING PEPTIDE.

Query Match 100.0%; Score 31; DB 1; Length 255;
 Best Local Similarity 100.0%; Pred. No. 25;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFPD 5
 |||||
 DB 78 GGFPD 82

RESULT 7
 YRFD_ECOLI
 ID YRFD_ECOLI STANDARD; PRT; 259 AA.
 AC P45753;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein yrfd.
 GN YRFD OR B3395.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 CC Escherichia coli.
 NCBI_Taxid=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12 / MG1655;
 RX MEDLINE-9742617; PubMed-9278503;
 RA Blatter F.R., Plunkett G., Ilt, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Zhao Y.;
 RA "The complete genome sequence of Escherichia coli K-12";
 RL Science 277:1453-1474(1997).
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U18997; AAA58192.1; ALT_INIT.
 DR EMBL; AE000414; AAC76420.1; ALT_INIT.
 DR Ecogene; EG12925; yrfd.
 KM Hypothetical protein; Complete proteome.
 SO SEQUENCE 259 AA; 29000 MW; B3E8B4ADE4AB16D CRC64;

Query Match 100.0%; Score 31; DB 1; Length 259;
 Best Local Similarity 100.0%; Pred. No. 26;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 GGFPD 5
 |||||
 DB 225 GGFPD 229

RESULT 8
 PSBP_CUCSA
 ID PSBP_CUCSA STANDARD; PRT; 263 AA.
 AC Q9SL08;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE Oxygen-evolving enhancer protein 2, chloroplast precursor (OEE2) (23
 DE kDa subunit of oxygen evolving system of photosystem II) (OEC 23 kDa
 DE subunit) (23 kDa thylakoid membrane protein) (OEC23).
 GN PSBP.
 OS Cucumis sativus (Cucumber).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosida I; Cucurbitales; Cucurbitaceae; Cucumis.
 NCBI_Taxid=3659;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-cv. Acnagajbaj;
 RA Ifuku K., Kanda Y., Sato F.;
 RT "Isolation of cucumber cDNA for the 23 kDa polypeptide of the oxygen-
 RT evolving complex of photosystem II (psbp).";
 RL (In) Plant Gene Register PGR00-031.
 CC -1- FUNCTION: MAY BE INVOLVED IN THE REGULATION OF PHOTOSYSTEM II.
 CC -1- SUBCELLULAR LOCATION: Chloroplast thylakoid membrane; associated
 CC with the photosystem II complex.
 CC -1- SIMILARITY: BELONGS TO THE PSBP FAMILY.
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AB032245; BAA69317.1; -
 DR InterPro; IPR002683; PSII_PsbP.
 DR Pfam; PF01789; PsbP; 1.
 KM Photosynthesis; Photosystem II; Chloroplast; Transl peptide;
 KM Thylakoid; Membrane.
 FT TRANSIT 1 77
 FT CHAIN 78 263
 FT SEQUENCE 263 AA; 28138 MW; 62EC7972B1254605 CRC64;
 OY 1 GGFPD 5
 |||||
 DB 178 GGFPD 182

Query Match 100.0%; Score 31; DB 1; Length 263;
 Best Local Similarity 100.0%; Pred. No. 26;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 GGFPD 5
 |||||
 DB 178 GGFPD 182

RESULT 9
 PSBP_FRING
 ID PSBP_FRING STANDARD; PRT; 264 AA.
 AC Q49080;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Oxygen-evolving enhancer protein 2, chloroplast precursor (OEE2) (23
 DE kDa subunit of oxygen evolving system of photosystem II) (OEC 23 kDa
 DE subunit) (23 kDa thylakoid membrane protein) (OEC23).
 GN PSBP.
 OS Cucumis sativus (Cucumber).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosida I; Cucurbitales; Cucurbitaceae; Cucumis.
 NCBI_Taxid=3659;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-cv. Acnagajbaj;
 RA Ifuku K., Kanda Y., Sato F.;
 RT "Isolation of cucumber cDNA for the 23 kDa polypeptide of the oxygen-
 RT evolving complex of photosystem II (psbp).";
 RL (In) Plant Gene Register PGR00-031.
 CC -1- FUNCTION: MAY BE INVOLVED IN THE REGULATION OF PHOTOSYSTEM II.
 CC -1- SUBCELLULAR LOCATION: Chloroplast thylakoid membrane; associated
 CC with the photosystem II complex.
 CC -1- SIMILARITY: BELONGS TO THE PSBP FAMILY.
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U18997; AAA58192.1; ALT_INIT.
 DR EMBL; AE000414; AAC76420.1; ALT_INIT.
 DR Ecogene; EG12925; yrfd.
 KM Hypothetical protein; Complete proteome.
 SO SEQUENCE 259 AA; 29000 MW; B3E8B4ADE4AB16D CRC64;

DE subunit) (23 kDa thylakoid membrane protein).
 GN PSBP
 OS Fritillaria agrestis.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Liliales; Liliaceae;
 OC Fritillaria.
 ON NCBI_TaxID=64177;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Paulco E., Baysdorfer C.;
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: MAY BE INVOLVED IN THE REGULATION OF PHOTOSYSTEM II.
 CC -1- SUBCELLULAR LOCATION: Chloroplast thylakoid membrane; associated
 CC with the photosystem II complex.
 CC -1- SIMILARITY: BELONGS TO THE PSBP FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF037458; AAC04809.1; -
 CC InterPro: IPR002683; PSII_PSBP.
 DR Pfam: PF01789; Psbp.1.
 KW Photosynthesis; Photosystem II; Chloroplast; Transit peptide;
 KM Thylakoid; Membrane.
 FT TRANSIT 1
 FT CHAIN 79 264 OXYGEN-EVOLVING ENHANCER PROTEIN 2.
 SQ SEQUENCE 264 AA; 28111 MW; 1E214D43C95809B CRC64;
 OY 1 GGFPD 5
 Db 179 GGFPD 183
 RESULT 10
 T4S9_HUMAN STANDARD; PRT; 268 AA.
 ID T4S9_HUMAN
 AC 060628; 060746; 09JULY1;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Transmembrane 4 superfamily, member 8 (Tetraspanin 5) (Tspan-5)
 DE (Tetraspan NET-4).
 GN TM4SF9 OR TSPAN5.
 OS Homo sapiens (Human), and
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 ON NCBI_TaxID=9606, 10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA SPECIES=Human;
 RC Rudlinshtein E., Serru V., Boucheix C.;
 RT New tetraspans identified in the EST database.
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 SQ SEQUENCE FROM N.A.

RC SPECIES=Human; TISSUE=Pancreas;
 RA Strausberg R.;
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Garcia-Frigola C., de Lecea L., Soriano E.;
 RT "Mouse tspan-5 cDNA cloning."
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
 CC -1- SIMILARITY: BELONGS TO THE TETRASPANIN (TM4SF) FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF053455; AAC69712.1; -
 CC EMBL: AF065389; AAC17120.1; -
 CC EMBL: BC009704; AAH09704.1; -
 CC EMBL: AF121344; AAF28869.1; -
 CC MGD: MGI:1928096; Tm4sf9.
 DR InterPro: IPR000301; Transmem_4.
 DR Pfam: PF00335; Transmembrane4; 1.
 DR PRINTS: PR00259; TMFOUR.
 DR PROSITE: PS00421; TM4_1; FALSE_NEG.
 KW Glycoprotein; Transmembrane.
 FT DOMAIN 1 17
 FT TRANSSEM 1 38
 FT DOMAIN 39 61
 FT TRANSSEM 62 82
 FT DOMAIN 83 92
 FT TRANSSEM 93 113
 FT DOMAIN 114 232
 FT TRANSSEM 233 253
 FT DOMAIN 254 268
 FT CARBOHYD 49 49
 FT CARBOHYD 169 169
 FT CARBOHYD 174 174
 FT CARBOHYD 232 232
 FT CONFLICT 91 94
 SQ SEQUENCE 268 AA; 30337 MW; 7F4480BD0FA6192D CRC64;
 OY 1 GGFPD 5
 Db 57 GGFPD 61
 RESULT 11
 FZF1_YEAST STANDARD; PRT; 299 AA.
 ID FZF1_YEAST
 AC P32805;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Zinc finger protein FZF1 (Sulphite resistance protein 1).
 GN FZF1 OR SUL1 OR YGL254W OR NCR299.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycaceae;
 ON NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA SRAIN-W303;
 RC MEDLINE-93311123; PubMed-8322518;
 RX Breitwieser W., Price C., Schuster T.;

```

RT *Identification of a gene encoding a novel zinc finger protein in
RT Saccharomyces cerevisiae.*;
RT Yeast 9:551-556(1993).
RN [2]
RC SEQUENCE FROM N.A.
RP STRAIN-S288C;
RX MEDLINE=95084637; PubMed=7992510;
RA Casalone E., Colella C.M., Daly S., Fontana S., Torricelli I.,
RA Polistrelli M.;
RT "Cloning and characterization of a sulphite-resistance gene of
RT Saccharomyces cerevisiae.";
RL Yeast 10:1101-1110(1994).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C / FY1679;
RX MEDLINE=97127827; PubMed=8972578;
RA Colisac E., Maillet E., Robineau S., Netter P.;
RT "Sequence of a 39,411 bp DNA fragment covering the left end of
RT chromosome VII of Saccharomyces cerevisiae.";
RL Yeast 12:1555-1562(1996).
CC -1- FUNCTION: MAY SERVE AS A TRANSCRIPTION FACTOR.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (POSSIBLE).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: X67787; CAA47998.1; -
DR EMBL: X78104; CAA54996.1; -
DR EMBL: X94357; CAA64133.1; -
DR EMBL: Z72776; CAA6974.1; -
DR PIR: S28554; S28554.
DR PIR: S33655; S33655.
DR HSSP: P03001; ITR3.
DR TRANSFAC: T03312; -
DR SGD: S0003223; FZF1.
DR InterPro: IPR000822; ZnF_C2H2.
DR Pfam: PF00096; zf_C2H2; 4.
DR PRINTS: PR00048; ZINCINGER.
DR SMART: SM00355; ZnF_C2H2; 5.
DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 3.
DR PROSITE: PS0157; ZINC_FINGER_C2H2_2; 3.
DR KW Transcription regulation; DNA-binding; Nuclear protein; Zinc-finger;
KW Metal-binding; Repeat.
FT ZN_FING 12 36 C2H2-TYPE.
FT ZN_FING 42 66 C2H2-TYPE.
FT ZN_FING 72 94 C2H2-TYPE.
FT DOMAIN 96 107 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT ZN_FING 155 179 C2H2-TYPE.
FT ZN_FING 246 271 C2H2-TYPE.
FT CONFLICT 180 180 H->D (IN REF. 2).
FT CONFLICT 235 235 T->W (IN REF. 1).
FT CONFLICT 246 246 N->Y (IN REF. 2).
SQ SEQUENCE 299 AA; 33994 MW; AAE7257F012B32FF CMC64;

Query Match 100.0%; Score 31; DB 1; Length 299;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFDP 5
DB 124 GGFDP 128

```

```

DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE N5, N10-methenyltetrahydromethanopterin cyclohydrolase (EC 3.5.4.27)
DE (Methenyl-H4MPT cyclohydrolase).
GN MCH.
OS Xanthobacter autotrophicus.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Hypnomicrobium group; Xanthobacter.
OX NCBI_TaxId=280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 35674 / DSM 432;
RX MEDLINE=99412275; PubMed=10482517;
RA Vorholt J.A., Chistoserdova L.V., Stoljar S.M., Thauer R.K.,
RA Lidstrom M.E.;
RT "Distribution of tetrahydromethanopterin-dependent enzymes in
RT methylophilic bacteria and phylogeny of methenyl
RT tetrahydromethanopterin cyclohydrolases.";
RL J. Bacteriol. 181:5750-5757(1999).
CC -1- FUNCTION: Catalyzes the hydrolysis of methenyl-H(4)MPT(+) to N(5)-
CC formyl-H(4)MPT (by similarity).
CC -1- CATALYTIC ACTIVITY: 5,10-methenyl-5,6,7,8-tetrahydromethanopterin
CC + H(2)O = N5-formyl-5,6,7,8-tetrahydromethanopterin.
CC -1- PATHWAY: H(4)MPT-dependent pathway of formaldehyde oxidation;
CC third step.
CC -1- SIMILARITY: BELONGS TO THE MCH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AF199593; AAD5896.1; -
DR HSSP: P94954; IOLM.
DR InterPro: IPR003209; Cyclohydrolase.
DR Pfam: PF02289; MCH; 1.
DR KW Hydrolase; One-carbon metabolism.
SQ SEQUENCE 337 AA; 34782 MW; AC53B27EBDDFBF08 CRC64;

Query Match 100.0%; Score 31; DB 1; Length 337;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFDP 5
DB 324 GGFDP 328

```

```

RESULT 13
DICH_DROME STANDARD; PRT; 382 AA.
AC Q24533; G9VUD4;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE SOX-domain protein dichaeete (Fish-hook protein).
GN D OR FISH OR SOX70D OR CG5893.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxId=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Oregon-R;
RX MEDLINE=97108761; PubMed=8951082;
RA Russell S.R.H., Sanchez-Soriano N., Wright C.R., Ashburner M.;
RT "The Dichaeete gene of Drosophila melanogaster encodes a SOX-domain
RT protein required for embryonic segmentation.";

```

Development 122:3669-3676(1996).

SEQUENCE FROM N.A.

MEDLINE-97108741; PubMed-8951062;

Nambu P.A., Nambu J.R.;

"The Drosophila fish-hook gene encodes a HMG domain protein essential for segmentation and CNS development.";

Development 122:3467-3475(1996).

SEQUENCE FROM N.A.

STRAIN-Berkeley;

MEDLINE-20196006; PubMed-10731132;

Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D., Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F., George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X., Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D., Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G., Abail J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D., Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M., Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S., Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotler P., Burks K.C., Busam D.A., Butler H., Cadiou E., Center A., Chandra I., Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P., de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M., Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P., Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W., Foeller C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K., Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M., Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J., Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C., Jalili M., Kalush F., Karen G.H., Ke Z., Kenton J.A., Ketchum K.A., Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z., Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X., Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D., Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A., Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L., Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M., Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G., Rhine K., Remington K., Sanders R.D.C., Scheefel F., Shen H., Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T., Spletter E., Spradling A.C., Stapleton M., Strong R., Sun E., Svirbas R., Tector R., Turner R., Venter E., Wang A.H., Wang X., Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J., Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L., Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O., Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

"The genome sequence of Drosophila melanogaster.";

Science 287:2185-2195(2000).

FUNCTION: ESSENTIAL FOR SEGMENTATION AND CNS DEVELOPMENT. MAY MODULATE THE ACTIONS OF OTHER TRANSCRIPTION FACTORS, INCLUDING GAP AND PAIR-RULE PROTEINS.

SUBCELLULAR LOCATION: Nucleus (Potential).

TISSUE SPECIFICITY: INITIALLY EXPRESSED IN A PAIR-RULE-LIKE PATTERN WHICH IS RAPIDLY REPLACED BY STRONG NEUROECTODERM EXPRESSION.

DEVELOPMENTAL STAGE: EXPRESSED DURING EMBRYOGENESIS.

SIMILARITY: CONTAINS 1 HMG BOX.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL: X96419; CA65279.1;

EMBL: U68056; AAB45673.1;

EMBL: AE003353; AAF49753.1;

HSSP: P48436; ISX9.

flyBase: FBgn0000411; D.

InterPro: IPR000910; HMG_12_box.

Pfam: PF00505; HMG_box; 1.

SMART: SM00398; HMG; 1.

Developmental protein; DNA-binding; Nuclear protein.

FT DNAS_BIND 142 210 HMG_BOX.

FT DOMAIN 224 251 GLY-RICH.

FT DOMAIN 293 298 POLY-ALA.

SO SEQUENCE 382 AA; 40212 MW; B03A267F9230A054 CRC64;

Query Match 100.0%; Score 31; DB 1; Length 382;

Best Local Similarity 100.0%; Pred. No. 38;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCFDP 5

Db 278 GCFDP 282

RESULT 14

ASSY_SCHPO	STANDARD	PRT	410 AA.
AC 094354;			
DT 30-MAY-2000 (Rel. 39, Created)			
DT 30-MAY-2000 (Rel. 39, Last sequence update)			
DT 15-JUN-2002 (Rel. 41, Last annotation update)			
DE Argininosuccinate synthase (EC 6.3.4.5) (Citruiline--aspartate ligase).			
GN SPCC428.05C.			
OS Schizosaccharomyces pombe (fission yeast).			
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;			
OC Schizosaccharomycetales; Schizosaccharomycetaceae;			
OX NCBI_TaxID=4896;			
OX NCBI_TaxID=4896;			
SEQUENCE FROM N.A.			
STRAIN-972;			
MEDLINE-21848401; PubMed-11859360;			
Wood V., Gwilliam R., Ratendream M.A., Lyne M., Lyne R., Stewart A., Sgourou J., Peat N., Hayles J., Baker S., Basham D., Bowman S., Brooks K., Brown D., Brown N., Chillingworth T., Churcher C.M., Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A., Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G., Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K., James K., Jones L., Jones M., Leather S., McDonald S., McLean J., Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odeli C., Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E., Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S., Skelton J., Simmonds M., Squares R., Squares S., Stevens K., Taylor K., Taylor R.G., Tiley A., Walsh S.V., Warren T., Whitehead S., Woodward J., Voicakeart G., Aert R., Robben J., Gymnopoulos B., Weljens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Filiz C., Holzer E., Moestl D., Hilbert H., Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M., Eger P., Zimmermann W., Wedler H., Wamutt R., Purnelle B., Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S., Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M., Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G., Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J., Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L., Cerretti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J., Sipakovski G.V., Ussery D., Bartell B.G., Nurse P.; <p>"The genome sequence of Schizosaccharomyces pombe.";</p> <p>Nature 415:871-880(2002).</p> <p>CATALYTIC ACTIVITY: ATP + L-citrulline + L-aspartate -> AMP + dihydrophosphate + L-argininosuccinate</p> <p>PATHWAY: UREA CYCLE, PENULTIMATE STEP OF THE ARGININE BIOSYNTHETIC PATHWAY.</p> <p>SUBUNIT: HOMOTETRAMER (BY SIMILARITY).</p> <p>SIMILARITY: BELONGS TO THE ARGININOSUCCINATE SYNTHASE FAMILY.</p> <p>This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its</p>			


```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AL034382; CAA22280.1; -
DR InterPro: IPR001518; Arginosuc_synth.
DR Pfam: PF00764; Arginosuc_synth; 1.
DR ProDom: PD003544; Arginosuc_synth; 1.
DR TIGRFAMs: TIGR00032; argG; 1.
DR PROSITE: PS00564; ARGININOSUCCIN_SYN_1; 1.
DR PROSITE: PS00565; ARGININOSUCCIN_SYN_2; 1.
DR Arginine biosynthesis; urea cycle; Ligase; ATP-binding.
FT NP_BIND 116 124 ATP (POTENTIAL).
SQ SEQUENCE 410 AA; 46084 MW; 1B3B873A4DCDC21 CRC64;

Query Match 100.0%; Score 31; DB 1; Length 410;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFDP 5
DB 382 GGFDP 386

RESULT 15
ID GLPB_ECOLI STANDARD; PRT; 419 AA.
AC P13033;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Anaerobic glycerol-3-phosphate dehydrogenase subunit B (EC 1.1.99.5)
DE (G-3-P dehydrogenase).
GN GLPB OR B2242.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-10.
RC STRAIN-K12;
RX MEDLINE=88227815; PubMed=3286606;
RA Cole S.T., Eljemeier K., Ahmed S., Honore N., Elmes L., Anderson W.F.,
RA Weiner J.H.;
RT "Nucleotide sequence and gene-polypeptide relationships of the glpABC
RT operon encoding the anaerobic sn-glycerol-3-phosphate dehydrogenase
RT of Escherichia coli K-12."
RL J. Bacteriol. 170:2448-2456(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-K12;
RX MEDLINE=97349980; PubMed=9205837;
RA Yamamoto Y., Alba H., Baba H., Hayashi K., Inada T., Isono K.,
RA Itoh T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsuhashi N.,
RA Mizobuchi K., Mori H., Nakade S., Nakamura Y., Nishimoto H.,
RA Oshima T., Oyama S., Saito N., Sempel G., Satoh Y., Sivasubraman S.,
RA Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,
RA Yamagata S., Horinouchi T.;
RT "Construction of a contiguous 874-kb sequence of the Escherichia coli
RT K-12 genome corresponding to 50.0-66.8 min on the linkage map and
RT DNA Res. 4:91-113(1997).

```

```

CC -1- FUNCTION: CONVERSION OF GLYCEROL 3-PHOSPHATE TO DIHYDROXYACETONE.
CC USES FUMARATE OR NITRATE AS ELECTRON ACCEPTOR.
CC -1- CATALYTIC ACTIVITY: sn-glycerol 3-phosphate + acceptor = glycerone
CC phosphate + reduced acceptor.
CC -1- COFACTOR: FMN IS SAID TO BE ASSOCIATED WITH GLPB.
CC -1- PATHWAY: FIRST STEP IN THE TRANSFER OF REDUCING EQUIVALENTS
CC FROM G-3-P TO A SHORT ELECTRON TRANSFER CHAIN TERMINATING WITH
CC FUMARATE AS THE ULTIMATE ELECTRON ACCEPTOR.
CC -1- SUBUNIT: COMPOSED OF A CATALYTIC GLPB/B DIMER AND OF MEMBRANE
CC BOUND GLPB.
CC -1- SUBCELLULAR LOCATION: LOOSELY BOUND TO THE CYTOPLASMIC MEMBRANE
CC OFTEN OCCURRING IN VESICLES ASSOCIATED WITH FUMARATE REDUCTASE.
CC -1- SIMILARITY: TO H. INFLUENZAE GLPB.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M20938; AAA83865.1; -
DR EMBL: AE000314; AAC75302.1; -
DR EMBL: D90855; BAA16061.1; -
DR PIR: B32006; DECBN.
DR BioGene: EG10392; glpB.
DR Oxidoreductase; Flavoprotein; FMN; Complete proteome.
SQ SEQUENCE 419 AA; 45357 MW; C8A2285AD09FAF55 CRC64;

Query Match 100.0%; Score 31; DB 1; Length 419;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFDP 5
DB 386 GGFDP 390

```

Search completed: March 17, 2003, 12:07:14
Job time : 11 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 17, 2003, 12:05:48 ; Search time 25 Seconds
(without alignments)
41.209 Million cell updates/sec

Title: SEQ25-FD
Perfect score: 31
Sequence: 1 gffdp 5

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP invertebrate:*
6: SP mammal:*
7: SP mhc:*
8: SP organelle:*
9: SP phage:*
10: SP plant:*
11: SP rodent:*
12: SP virus:*
13: SP vertebrate:*
14: SP unclassified:*
15: SP viirus:*
16: SP bacteriap:*
17: SP archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31	100.0	56	7 098048	098048 monachus sc
2	31	100.0	61	6 019142	019142 cercopithec
3	31	100.0	61	6 019149	019149 papio anubi
4	31	100.0	61	6 019150	019150 papio anubi
5	31	100.0	61	6 019151	019151 papio anubi
6	31	100.0	61	6 019159	019159 cercopithec
7	31	100.0	61	6 019161	019161 cercopithec
8	31	100.0	62	6 019140	019140 theropithec
9	31	100.0	62	6 019146	019146 cercopithec
10	31	100.0	62	6 019153	019153 papio anubi
11	31	100.0	62	6 019157	019157 cercopithec
12	31	100.0	62	6 019160	019160 cercopithec
13	31	100.0	62	6 019163	019163 cercopithec
14	31	100.0	62	6 019170	019170 cercopithec
15	31	100.0	62	7 019300	019300 gorilla gor
16	31	100.0	62	7 098035	098035 macaca mula

17	31	100.0	62	7 098037	098037 macaca mula
18	31	100.0	62	7 098039	098039 macaca mula
19	31	100.0	62	7 098141	098141 papio cynoc
20	31	100.0	62	7 098147	098147 papio cynoc
21	31	100.0	62	7 098148	098148 papio cynoc
22	31	100.0	62	7 098150	098150 papio cynoc
23	31	100.0	62	7 098151	098151 papio hamad
24	31	100.0	62	7 030994	030994 pan troglod
25	31	100.0	62	7 030684	030684 macaca mula
26	31	100.0	62	7 030685	030685 macaca mula
27	31	100.0	62	7 030552	030552 macaca arc
28	31	100.0	62	7 030553	030553 macaca arc
29	31	100.0	62	7 030575	030575 macaca arc
30	31	100.0	62	7 030531	030531 macaca fasc
31	31	100.0	62	7 030902	030902 pongo pygma
32	31	100.0	62	7 030906	030906 pongo pygma
33	31	100.0	62	7 030689	030689 macaca mula
34	31	100.0	63	6 019143	019143 cercopithec
35	31	100.0	63	6 019144	019144 cercopithec
36	31	100.0	63	6 019148	019148 papio cynoc
37	31	100.0	63	7 078025	078025 callithrix
38	31	100.0	63	7 098142	098142 papio cynoc
39	31	100.0	63	7 098143	098143 papio cynoc
40	31	100.0	63	7 030336	030336 cercopithec
41	31	100.0	63	7 030681	030681 macaca mula
42	31	100.0	63	7 030683	030683 macaca mula
43	31	100.0	63	7 030576	030576 macaca fasc
44	31	100.0	63	7 030690	030690 macaca mula
45	31	100.0	70	7 097NR8	097NR8 homo sapien

ALIGNMENTS

RESULT 1

ID 098048 PRELIMINARY; PRT: 56 AA.
AC 098048:
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE MHC class II antigen (Fragment).
DOA.
OS Monachus schauinslandi (Hawaiian monk seal).
OC Eukaryota; Metazoa; Chordata; Cranata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Plinipedia; Phocidae; Monachus.
OX NCBI_TaxID=29088;
RN [1]
RP SEQUENCE FROM N.A.
RA Armstrong P.S.;
RT "Hawaiian monk seal DOA gene exon 2."
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF093799; AAC83221.1;
DR InterPro; IPR001003; MHC_II_alpha.
DR Pfam; PF009933; MHC_II_alpha; 1.
FT NON_TER 1
FT NON_TER 56
SQ SEQUENCE 56 AA; 6409 MW; E0738E707EC74B33 CRC64;

Query Match 100.0%; Score 31; DB 7; Length 56;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGPP 5
DB 24 GGPP 28

RESULT 2

ID 019142 PRELIMINARY; PRT: 61 AA.
AC 019142:
DT 01-JAN-1998 (TREMBlrel. 05, Created)

DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DE 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 AC Mhc-DQAI (Fragment).
 OS Cercopithecus aethiops.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecinae; Cercopithecidae;
 DE Mhc-DQAI (Fragment).
 RN NCBI_TaxID=36222;
 RP SEQUENCE FROM N.A.
 RC STRAIN=BLM29;
 RX MEDLINE=97364700; PubMed=9218546;
 RA Mwendu J.M., Takenaka O., Kim H.-S., Yamamoto T., Katsumata Y.,
 RA Katsumata Y., Bamba C.S., Uchihl R., Shotake T.;
 RT "Nucleotide sequences of the major histocompatibility complex DQAI
 RT locus of Cercopithecus monkeys.";
 RL Immunogenetics 46:363-364(1997).
 DR EMBL; D88643; BAA22110.1; -;
 DR InterPro: IPR001003; MHC_II_alpha.
 DR Pfam: PF00993; MHC_II_alpha; 1.
 FT NON_TER 1 61
 FT SEQUENCE 61 AA; 7196 MW; 035E3AB6DA63161E CRC64;

Query Match
 Best Local Similarity 100.0%; Score 31; DB 6; Length 61;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFPD 5
 DB 38 GGFPD 42

RESULT 3
 ID 019149 PRELIMINARY; PRT; 61 AA.
 AC 019149;
 DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Mhc-DQAI (Fragment).
 OS Papio anubis (Olive baboon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecinae; Cercopithecidae;
 DE NCBI_TaxID=9555;
 RN NCBI_TaxID=9555;
 RP SEQUENCE FROM N.A.
 RC STRAIN=PAN 1979;
 RX MEDLINE=97364700; PubMed=9218546;
 RA Mwendu J.M., Takenaka O., Kim H.-S., Yamamoto T., Gurja B.,
 RA Katsumata Y., Bamba C.S., Uchihl R., Shotake T.;
 RT "Major histocompatibility complex DQAI nucleotide sequences of gelada
 RT baboon (Theropithecus gelada), olive baboon (Papio anubis), and yellow
 RT baboon (Papio cynocephalus).";
 RL Immunogenetics 46:365-366(1997).
 DR EMBL; D88643; BAA22110.1; -;
 DR InterPro: IPR001003; MHC_II_alpha.
 DR Pfam: PF00993; MHC_II_alpha; 1.
 FT NON_TER 1 61
 FT SEQUENCE 61 AA; 7180 MW; EC413AB6DA741D73 CRC64;

Query Match
 Best Local Similarity 100.0%; Score 31; DB 6; Length 61;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFPD 5
 DB 38 GGFPD 42

RESULT 4

019150
 ID 019150 PRELIMINARY; PRT; 61 AA.
 AC 019150;
 DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Mhc-DQAI (Fragment).
 RN NCBI_TaxID=9555;
 RP SEQUENCE FROM N.A.
 RC STRAIN=PAN 2022;
 RX MEDLINE=97364700; PubMed=9218546;
 RA Mwendu J.M., Takenaka O., Kim H.-S., Yamamoto T., Gurja B.,
 RA Katsumata Y., Bamba C.S., Uchihl R., Shotake T.;
 RT "Major histocompatibility complex DQAI nucleotide sequences of gelada
 RT baboon (Theropithecus gelada), olive baboon (Papio anubis), and yellow
 RT baboon (Papio cynocephalus).";
 RL Immunogenetics 46:365-366(1997).
 DR EMBL; D88644; BAA22111.1; -;
 DR InterPro: IPR001003; MHC_II_alpha.
 DR Pfam: PF00993; MHC_II_alpha; 1.
 FT NON_TER 1 61
 FT SEQUENCE 61 AA; 7180 MW; EC413AB6DA741D73 CRC64;

Query Match
 Best Local Similarity 100.0%; Score 31; DB 6; Length 61;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFPD 5
 DB 38 GGFPD 42

RESULT 5
 ID 019151 PRELIMINARY; PRT; 61 AA.
 AC 019151;
 DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Mhc-DQAI (Fragment).
 OS Papio anubis (Olive baboon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecinae; Cercopithecidae;
 DE NCBI_TaxID=9555;
 RN NCBI_TaxID=9555;
 RP SEQUENCE FROM N.A.
 RC STRAIN=PAN 2041;
 RX MEDLINE=97364700; PubMed=9218546;
 RA Mwendu J.M., Takenaka O., Kim H.-S., Yamamoto T., Gurja B.,
 RA Katsumata Y., Bamba C.S., Uchihl R., Shotake T.;
 RT "Major histocompatibility complex DQAI nucleotide sequences of gelada
 RT baboon (Theropithecus gelada), olive baboon (Papio anubis), and yellow
 RT baboon (Papio cynocephalus).";
 RL Immunogenetics 46:365-366(1997).
 DR EMBL; D88645; BAA22112.1; -;
 DR InterPro: IPR001003; MHC_II_alpha.
 DR Pfam: PF00993; MHC_II_alpha; 1.
 FT NON_TER 1 61
 FT SEQUENCE 61 AA; 7180 MW; EC413AB6DA741D73 CRC64;

Query Match
 Best Local Similarity 100.0%; Score 31; DB 6; Length 61;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFPD 5

Db 38 GGFDP 42

RESULT 6
ID 019159 PRELIMINARY; PRT; 61 AA.

AC 019159;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Mhc-DQAI (Fragment).
OS Cercopithecus mltis.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Cercopithecus.
NCBI_TaxID=36225;

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SVK354;
RA MEDLINE-97364699; PubMed-9218545;
RA Mwenda J.M., Takenaka O., Kim H.-S., Yamamoto T., Katsumata Y.,
Bambra C.S., Uchihl R., Shotake T.,
RT "Nucleotide sequences of the major histocompatibility complex DQAI
locus of Cercopithecus monkeys.";
RL Immunogenetics 46:363-634(1997).
DR EMBL; D88680; BAA22120.1; -
DR InterPro: IPR001003; MHC_II_alpha.
DR Pfam: PF00993; MHC_II_alpha; 1.
FT NON_TER 1
FT SEQUENCE 61 AA; 7230 MM; 667A28157A740393 CRC64;

Query Match 100.0%; Score 31; DB 6; Length 61;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFDP 5
Db 38 GGFDP 42

RESULT 7
ID 019161 PRELIMINARY; PRT; 61 AA.

AC 019161;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Mhc-DQAI (Fragment).
OS Cercopithecus mltis.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Cercopithecus.
NCBI_TaxID=36225;

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SYL70;
RA MEDLINE-97364699; PubMed-9218545;
RA Mwenda J.M., Takenaka O., Kim H.-S., Yamamoto T., Katsumata Y.,
Bambra C.S., Uchihl R., Shotake T.,
RT "Nucleotide sequences of the major histocompatibility complex DQAI
locus of Cercopithecus monkeys.";
RL Immunogenetics 46:363-634(1997).
DR EMBL; D88682; BAA22122.1; -
DR InterPro: IPR001003; MHC_II_alpha.
DR Pfam: PF00993; MHC_II_alpha; 1.
FT NON_TER 1
FT SEQUENCE 61 AA; 7230 MM; 667A28157A740393 CRC64;

Query Match 100.0%; Score 31; DB 6; Length 61;
Best Local Similarity 100.0%; Pred. No. 52;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 GGFDP 5
Db 38 GGFDP 42

RESULT 8
ID 019140 PRELIMINARY; PRT; 62 AA.

AC 019140;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-DIC-2001 (TREMBlrel. 19, Last annotation update)
DE Mhc-DQAI (Fragment).
OS Theropithecus gelada (Gelada baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Theropithecus.
NCBI_TaxID=9565;

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A-13;
RA MEDLINE-97364700; PubMed-9218546;
RA Mwenda J.M., Takenaka O., Kim H.-S., Yamamoto T., Gurja B.,
Bambra C.S., Uchihl R., Shotake T.,
RT "Major histocompatibility complex DQAI nucleotide sequences of gelada
baboon (Theropithecus gelada), olive baboon (Papio anubis), and yellow
baboon (Papio cynocephalus)."
RL Immunogenetics 46:365-366(1997).
DR EMBL; D88550; BAA22101.1; -
DR InterPro: IPR001003; MHC_II_alpha.
DR Pfam: PF00993; MHC_II_alpha; 1.
FT NON_TER 1
FT SEQUENCE 62 AA; 7304 MM; 2C32654A733B1CEF CRC64;

Query Match 100.0%; Score 31; DB 6; Length 62;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFDP 5
Db 38 GGFDP 42

RESULT 9
ID 019146 PRELIMINARY; PRT; 62 AA.

AC 019146;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Mhc-DQAI (Fragment).
OS Cercopithecus aethiops (Green monkey) (Girvet).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Cercopithecus.
NCBI_TaxID=9534;

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-VER1621;
RA MEDLINE-97364699; PubMed-9218545;
RA Mwenda J.M., Takenaka O., Kim H.-S., Yamamoto T., Katsumata Y.,
Bambra C.S., Uchihl R., Shotake T.,
RT "Nucleotide sequences of the major histocompatibility complex DQAI
locus of Cercopithecus monkeys.";
RL Immunogenetics 46:363-634(1997).
DR EMBL; D88583; BAA22107.1; -
DR InterPro: IPR001003; MHC_II_alpha.
DR Pfam: PF00993; MHC_II_alpha; 1.
FT NON_TER 1
FT SEQUENCE 62 AA; 7304 MM; 2C32654A733B1CEF CRC64;

Query Match 100.0%; Score 31; DB 6; Length 62;
Best Local Similarity 100.0%; Pred. No. 53;

SO SEQUENCE 62 AA; 7105 MW; D49D54E4692010E0 CRC64;
 Query Match 100.0%; Score 31; DB 6; Length 62;
 Best Local Similarity 100.0%; Pred. No. 53;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFDP 5
 |||||
 Db 38 GGFDP 42

RESULT 10

ID 019153 PRELIMINARY; PRT; 62 AA.
 AC 019153;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Mhc-DQAI (Fragment).
 OS Papio anubis (Olive baboon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecinae; Papio.
 OX NCBI_TaxID=9555;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PAN 1991;
 RX MEDLINE=97364700; PubMed=9218546;
 RA Mwenda J.M., Takenaka O., Kim H.-S., Yamamoto T., Gurja B.,
 RA Katsumata Y., Bambara C.S., Uchihl R., Shotake T.,
 RT "Major histocompatibility complex DQAI nucleotide sequences of gelada
 baboon (Theropithecus gelada), olive baboon (Papio anubis), and yellow
 baboon (Papio cynocephalus).";
 RL Immunogenetics 46:363-366(1997).
 DR EMBL; D88647; BAA22114.1; -;
 DR InterPro: IPR001003; MHC_II_alpha.
 DR Pfam: PF00993; MHC_II_alpha; 1.
 FT NON_TER
 FT NON_TER
 SQ SEQUENCE 62 AA; 7214 MW; 66827EE1333110E4 CRC64;

Query Match 100.0%; Score 31; DB 6; Length 62;
 Best Local Similarity 100.0%; Pred. No. 53;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFDP 5
 |||||
 Db 38 GGFDP 42

RESULT 11

ID 019157 PRELIMINARY; PRT; 62 AA.
 AC 019157;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE Mhc-DQAI (Fragment).
 OS Cercopithecus mitis.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecinae; Cercopithecus.
 OX NCBI_TaxID=36225;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SVK269;
 RX MEDLINE=97364699; PubMed=9218545;
 RA Mwenda J.M., Takenaka O., Kim H.-S., Yamamoto T., Katsumata Y.,
 RA Bambara C.S., Uchihl R., Shotake T.,
 RT "Nucleotide sequences of the major histocompatibility complex DQAI
 locus of Cercopithecus monkeys";
 RL Immunogenetics 46:363-364(1997).
 DR EMBL; D88678; BAA22118.1; -;

Query Match 100.0%; Score 31; DB 6; Length 62;
 Best Local Similarity 100.0%; Pred. No. 53;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DR InterPro: IPR001003; MHC_II_alpha.
 DR Pfam: PF00993; MHC_II_alpha; 1.
 FT NON_TER
 FT NON_TER
 SQ SEQUENCE 62 AA; 7291 MW; AEB85E5E23393CD3 CRC64;

Query Match 100.0%; Score 31; DB 6; Length 62;
 Best Local Similarity 100.0%; Pred. No. 53;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFDP 5
 |||||
 Db 38 GGFDP 42

RESULT 12

ID 019160 PRELIMINARY; PRT; 62 AA.
 AC 019160;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE Mhc-DQAI (Fragment).
 OS Cercopithecus mitis.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecinae; Cercopithecus.
 OX NCBI_TaxID=36225;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SVL51;
 RX MEDLINE=97364699; PubMed=9218545;
 RA Mwenda J.M., Takenaka O., Kim H.-S., Yamamoto T., Katsumata Y.,
 RA Bambara C.S., Uchihl R., Shotake T.,
 RT "Nucleotide sequences of the major histocompatibility complex DQAI
 locus of Cercopithecus monkeys";
 RL Immunogenetics 46:363-364(1997).
 DR EMBL; D88681; BAA22121.1; -;
 DR InterPro: IPR001003; MHC_II_alpha.
 DR Pfam: PF00993; MHC_II_alpha; 1.
 FT NON_TER
 FT NON_TER
 SQ SEQUENCE 62 AA; 7394 MW; AF73F43E2331B1CF1 CRC64;

Query Match 100.0%; Score 31; DB 6; Length 62;
 Best Local Similarity 100.0%; Pred. No. 53;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFDP 5
 |||||
 Db 38 GGFDP 42

RESULT 13

ID 019163 PRELIMINARY; PRT; 62 AA.
 AC 019163;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE Mhc-DQAI (Fragment).
 OS Cercopithecus mitis.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecinae; Cercopithecus.
 OX NCBI_TaxID=36225;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SVL120;
 RX MEDLINE=97364699; PubMed=9218545;
 RA Mwenda J.M., Takenaka O., Kim H.-S., Yamamoto T., Katsumata Y.,
 RA Bambara C.S., Uchihl R., Shotake T.,
 RT "Nucleotide sequences of the major histocompatibility complex DQAI
 locus of Cercopithecus monkeys";
 RL Immunogenetics 46:363-364(1997).
 DR EMBL; D88678; BAA22118.1; -;

Query Match 100.0%; Score 31; DB 6; Length 62;
 Best Local Similarity 100.0%; Pred. No. 53;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

RT Locus of Cercopithecus monkeys."
RL Immunogenetics 46:363-634(1997).
EMBL: D88684: BAA22124.1.
DR InterPro: IPR001003; MHC_II_alpha.
DR Pfam: PF00993; MHC_II_alpha; 1.
FT NON_TER 1 1
SQ SEQUENCE 62 AA; 7381 MW; AF632E5E233B1CF1 CRC64;

Query Match
Best Local Similarity 100.0%; Score 31; DB 6; Length 62;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFDP 5
DB 38 GGFDP 42

RESULT 14
O19170 PRELIMINARY; PRT; 62 AA.
AC O19170:
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Mhc-DOAI (Fragment).
OS Cercopithecus mitis.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Cercopithecus.
ON NCBI_TaxID=36225;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SYLL19;
RX MEDLINE=97364699; PubMed=9218545;
RA Mwenda J.M., Takenaka O., Kim H.-S., Yamamoto T., Katsumata Y.,
RA Bamba C.S., Ochli R., Shokake T.;
RT "Nucleotide sequences of the major histocompatibility complex DOAI
RT locus of Cercopithecus monkeys."
RL Immunogenetics 46:363-634(1997).
EMBL: D89534: BAA22131.1.
DR InterPro: IPR001003; MHC_II_alpha.
DR Pfam: PF00993; MHC_II_alpha; 1.
FT NON_TER 1 1
SQ SEQUENCE 62 AA; 7374 MW; 8648A5E233B1B70 CRC64;

Query Match
Best Local Similarity 100.0%; Score 31; DB 6; Length 62;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFDP 5
DB 38 GGFDP 42

RESULT 15
O19300 PRELIMINARY; PRT; 62 AA.
AC O19300:
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE MHC class II DQ alpha (Fragment).
GN HLA-DOAI.
OS Gorilla gorilla (gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Gorilla.
ON NCBI_TaxID=9593;
RN [1]
RP SEQUENCE OF 6-27 FROM N.A.
RC STRAIN-NILMA; TISSUE-BLOOD;
RX MEDLINE=93339935; PubMed=8340229;

```

```

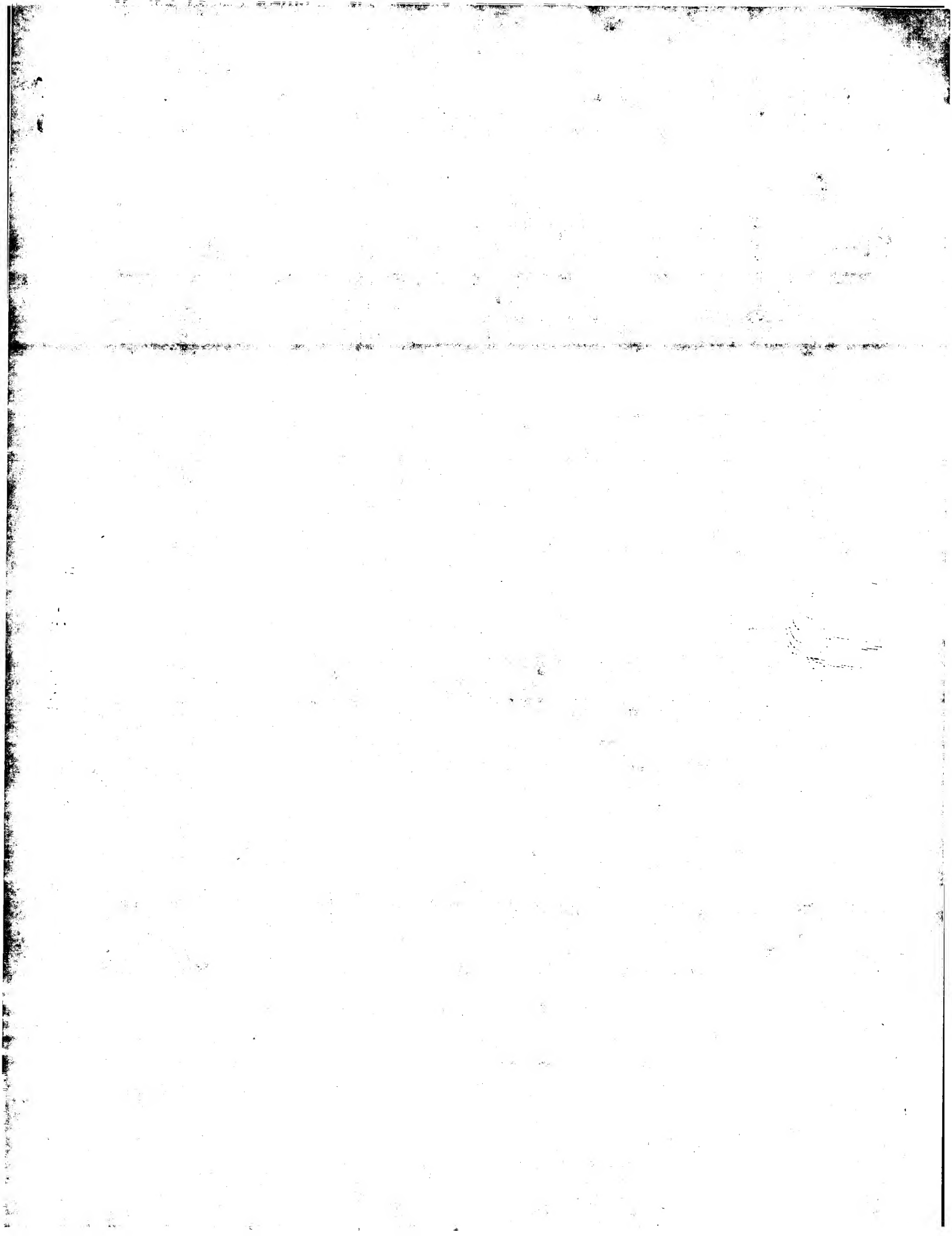
RA Kenter M., Otting N., de Weers M., Anholt J., Reiter C., Jonker M.,
RA Bontrop R.E.;
RT "Mhc-DRB and -DOAI nucleotide sequences of three lowland gorillas.
RT Implications for the evolution of primate Mhc class II haplotypes."
RL Hum. Immunol. 36:205-218(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-NILMA; TISSUE-BLOOD;
RA de Groot N.N.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF031277; AAB87448.1.
DR InterPro: IPR001003; MHC_II_alpha.
DR Pfam: PF00993; MHC_II_alpha; 1.
FT NON_TER 1 1
SQ SEQUENCE 62 AA; 7238 MW; 283A0016E981E8EF CRC64;

Query Match
Best Local Similarity 100.0%; Score 31; DB 7; Length 62;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFDP 5
DB 38 GGFDP 42

```

Search completed: March 17, 2003, 12:10:18
 Job time : 26 secs



GenCore version 5.1.4-p5-4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 17, 2003, 12:07:24 ; Search time 152 Seconds
(Without alignments)
21.208 Million cell updates/sec

Title: SEQ25-FD
Perfect score: 31
Sequence: 1 gqfdp 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_AA_Main:*

- 1: /cgn2_6/ptodata/2/paa/PCURS_COMB.pep:*
- 2: /cgn2_6/ptodata/2/paa/US06_COMB.pep:*
- 3: /cgn2_6/ptodata/2/paa/US07_COMB.pep:*
- 4: /cgn2_6/ptodata/2/paa/US080_COMB.pep:*
- 5: /cgn2_6/ptodata/2/paa/US081_COMB.pep:*
- 6: /cgn2_6/ptodata/2/paa/US082_COMB.pep:*
- 7: /cgn2_6/ptodata/2/paa/US083_COMB.pep:*
- 8: /cgn2_6/ptodata/2/paa/US084_COMB.pep:*
- 9: /cgn2_6/ptodata/2/paa/US085_COMB.pep:*
- 10: /cgn2_6/ptodata/2/paa/US086_COMB.pep:*
- 11: /cgn2_6/ptodata/2/paa/US087_COMB.pep:*
- 12: /cgn2_6/ptodata/2/paa/US088_COMB.pep:*
- 13: /cgn2_6/ptodata/2/paa/US089_COMB.pep:*
- 14: /cgn2_6/ptodata/2/paa/US090_COMB.pep:*
- 15: /cgn2_6/ptodata/2/paa/US091_COMB.pep:*
- 16: /cgn2_6/ptodata/2/paa/US092_COMB.pep:*
- 17: /cgn2_6/ptodata/2/paa/US093_COMB.pep:*
- 18: /cgn2_6/ptodata/2/paa/US094_COMB.pep:*
- 19: /cgn2_6/ptodata/2/paa/US095_COMB.pep:*
- 20: /cgn2_6/ptodata/2/paa/US096_COMB.pep:*
- 21: /cgn2_6/ptodata/2/paa/US097_COMB.pep:*
- 22: /cgn2_6/ptodata/2/paa/US098_COMB.pep:*
- 23: /cgn2_6/ptodata/2/paa/US099_COMB.pep:*
- 24: /cgn2_6/ptodata/2/paa/US100_COMB.pep:*
- 25: /cgn2_6/ptodata/2/paa/US101_COMB.pep:*
- 26: /cgn2_6/ptodata/2/paa/US102_COMB.pep:*
- 27: /cgn2_6/ptodata/2/paa/US60_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	31	100.0	11	1	PCT-US01-19110-3191
2	31	100.0	11	22	US-09-880-748-3191
3	31	100.0	38	21	US-09-708-427-48538
4	31	100.0	49	27	US-60-181-425-26
5	31	100.0	56	22	US-09-855-754-9
6	31	100.0	60	14	US-09-069-726-181

7	31	100.0	60	17	US-09-312-283A-181	Sequence 181, App
8	31	100.0	60	17	US-09-312-283A-320	Sequence 320, App
9	31	100.0	60	17	US-09-312-283B-181	Sequence 181, App
10	31	100.0	60	17	US-09-312-283B-320	Sequence 320, App
11	31	100.0	60	17	US-09-312-283C-181	Sequence 181, App
12	31	100.0	60	17	US-09-312-283C-320	Sequence 320, App
13	31	100.0	60	22	US-09-866-050A-181	Sequence 320, App
14	31	100.0	60	22	US-09-866-050A-320	Sequence 181, App
15	31	100.0	70	21	US-09-866-050A-1362	Sequence 1362, App
16	31	100.0	73	21	US-09-791-537-134706	Sequence 134706, App
17	31	100.0	73	26	US-10-235-954-1362	Sequence 1362, App
18	31	100.0	83	1	PCT-US00-07661-107	Sequence 107, App
19	31	100.0	83	1	PCT-US00-07661-108	Sequence 108, App
20	31	100.0	83	1	PCT-US00-07661-115	Sequence 115, App
21	31	100.0	83	1	PCT-US00-07661-116	Sequence 116, App
22	31	100.0	83	1	PCT-US00-07661-107	Sequence 107, App
23	31	100.0	83	1	PCT-US00-07661-108	Sequence 108, App
24	31	100.0	83	1	PCT-US00-07661-115	Sequence 115, App
25	31	100.0	83	1	PCT-US00-07661-116	Sequence 116, App
26	31	100.0	83	27	US-60-163-062-670	Sequence 670, App
27	31	100.0	84	27	US-60-163-062-853	Sequence 853, App
28	31	100.0	86	21	US-09-791-537-24530	Sequence 24530, A
29	31	100.0	87	21	US-09-791-537-97065	Sequence 97065, A
30	31	100.0	88	27	US-60-196-718-4591	Sequence 4591, App
31	31	100.0	91	18	US-09-450-969-6033	Sequence 6033, App
32	31	100.0	93	12	US-08-827-356-4177	Sequence 4177, App
33	31	100.0	93	20	US-09-611-529-5791	Sequence 5791, App
34	31	100.0	93	13	US-08-999-994-3	Sequence 3, App11
35	31	100.0	95	13	US-08-999-994-3	Sequence 80387, A
36	31	100.0	96	21	US-09-791-537-80387	Sequence 3519, App
37	31	100.0	97	21	US-09-758-671-3519	Sequence 3519, App
38	31	100.0	97	26	US-10-235-953-3519	Sequence 32, App1
39	31	100.0	98	21	US-09-751-181-32	Sequence 78163, A
40	31	100.0	98	21	US-09-791-537-78163	Sequence 114, App
41	31	100.0	99	17	PCT-US99-09970-114	Sequence 114, App
42	31	100.0	99	12	US-09-306-111-114	Sequence 25092, A
43	31	100.0	99	22	US-09-834-366-25092	Sequence 1, App11
44	31	100.0	99	26	US-10-220-382-1	Sequence 1, App11
45	31	100.0	99	27	US-60-186-854-1	

ALIGNMENTS

RESULT 1
PCT-US01-19110-3191
Sequence 3191, Application PC/TUS0119110
GENERAL INFORMATION:
APPLICANT: Human Genome Sciences, Inc.
TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
FILE REFERENCE: PE523PCT
CURRENT APPLICATION NUMBER: PCT/US01/19110
CURRENT FILING DATE: 2001-06-15
PRIOR APPLICATION NUMBER: 60/212,210
PRIOR FILING DATE: 2000-06-15
PRIOR APPLICATION NUMBER: 60/240,816
PRIOR FILING DATE: 2000-10-17
PRIOR APPLICATION NUMBER: 60/276,248
PRIOR FILING DATE: 2001-03-16
PRIOR APPLICATION NUMBER: 60/277,379
PRIOR FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/293,499
PRIOR FILING DATE: 2001-05-25
NUMBER OF SEQ ID NOS: 3239
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 3191
LENGTH: 11
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US01-19110-3191
Query Match 100.0%; Score 31; DB 1; Length 11;
Best Local Similarity 100.0%; Pred. No. 36;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGFPD 5
 Db 7 GGFPD 11

RESULT 2
 US-09-880-748-3191
 ; Sequence 3191, Application US/09880748
 ; GENERAL INFORMATION:
 ; APPLICANT: Ruben et al.
 ; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
 ; FILE REFERENCE: P5523
 ; CURRENT APPLICATION NUMBER: US/09/880,748
 ; CURRENT FILING DATE: 2001-06-15
 ; PRIOR APPLICATION NUMBER: 60/212,210
 ; PRIOR FILING DATE: 2000-06-15
 ; PRIOR APPLICATION NUMBER: 60/240,816
 ; PRIOR FILING DATE: 2000-10-17
 ; PRIOR APPLICATION NUMBER: 60/276,248
 ; PRIOR FILING DATE: 2001-03-16
 ; PRIOR APPLICATION NUMBER: 60/277,379
 ; PRIOR FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: 60/293,499
 ; PRIOR FILING DATE: 2001-05-25
 ; NUMBER OF SEQ ID NOS: 3239
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 3191
 ; LENGTH: 11
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-880-748-3191

Query Match
 Best Local Similarity 100.0%; Score 31; DB 22; Length 11;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGFPD 5
 Db 7 GGFPD 11

RESULT 3
 US-09-708-427-48538
 ; Sequence 48538, Application US/09708427
 ; GENERAL INFORMATION:
 ; APPLICANT: N. ALEXANDROV et al.
 ; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
 ; FILE REFERENCE: 2750-1243P
 ; CURRENT APPLICATION NUMBER: US/09/708,427
 ; CURRENT FILING DATE: 2000-11-09
 ; NUMBER OF SEQ ID NOS: 85364
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 48538
 ; LENGTH: 38
 ; TYPE: PRT
 ; ORGANISM: Zea mays subsp. mays
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: 1..38
 ; OTHER INFORMATION: Xaa is any amino acid
 ; NAME/KEY: misc_feature
 ; LOCATION: 1..38
 ; OTHER INFORMATION: Ceres Seq. ID 1855288
 US-09-708-427-48538

Query Match
 Best Local Similarity 100.0%; Score 31; DB 21; Length 38;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGFPD 5

Db 9 GGFPD 13

RESULT 4
 US-60-181-425-26
 ; Sequence 26, Application US/60181425
 ; GENERAL INFORMATION:
 ; APPLICANT: Bonazzi, Vivien
 ; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
 ; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
 ; FILE REFERENCE: C1000215
 ; CURRENT APPLICATION NUMBER: US/60/181,425
 ; CURRENT FILING DATE: 2000-02-09
 ; NUMBER OF SEQ ID NOS: 36
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 26
 ; LENGTH: 49
 ; TYPE: PRT
 ; ORGANISM: HUMAN
 US-60-181-425-26

Query Match
 Best Local Similarity 100.0%; Score 31; DB 27; Length 49;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGFPD 5
 Db 13 GGFPD 17

RESULT 5
 US-09-855-754-9
 ; Sequence 9, Application US/09855754
 ; GENERAL INFORMATION:
 ; APPLICANT: BOURSIAUX-EUDE, CAROLINE
 ; TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED
 ; TITLE OF INVENTION: REGIONS OF PERTACTIN IN BORDETELLA PERTUSSIS,
 ; TITLE OF INVENTION: BORDETELLA PERTUSSIS, AND BORDETELLA
 ; TITLE OF INVENTION: BRONCHISEPTICA, THEIR USE IN DIAGNOSTICS, AND IN
 ; FILE REFERENCE: 03495-0206-00000
 ; CURRENT APPLICATION NUMBER: US/09/855,754
 ; CURRENT FILING DATE: 2001-09-10
 ; PRIOR APPLICATION NUMBER: 60/206,969
 ; PRIOR FILING DATE: 2000-05-25
 ; NUMBER OF SEQ ID NOS: 24
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 9
 ; LENGTH: 56
 ; TYPE: PRT
 ; ORGANISM: Bordetella bronchiseptica
 US-09-855-754-9

Query Match
 Best Local Similarity 100.0%; Score 31; DB 22; Length 56;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGFPD 5
 Db 23 GGFPD 27

RESULT 6
 US-09-069-726-181
 ; Sequence 181, Application US/09069726
 ; GENERAL INFORMATION:
 ; APPLICANT: WATSON, JAMES D.
 ; TITLE OF INVENTION: COMPOSITIONS ISOLATED FROM
 ; TITLE OF INVENTION: SKIN CELLS AND METHODS FOR THEIR USE
 ; NUMBER OF SEQUENCES: 197

CORRESPONDENCE ADDRESS:

ADDRESSEE: Law Offices of Ann W. Speckman
STREET: 2601 Elliott Avenue, Suite 4185
CITY: Seattle
STATE: WA
COUNTRY: USA

ZIP: 98121
COUNTRY: USA

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/069,726

FILING DATE: 29-APR-1998

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

ATTORNEY/AGENT INFORMATION:

NAME: Sleeth, Janet

REGISTRATION NUMBER: 37,007

TELECOMMUNICATION INFORMATION:

TELEPHONE: 206-269-0565

TELEFAX: 206-269-0563

TELEX:

INFORMATION FOR SEQ ID NO: 181:

SEQUENCE CHARACTERISTICS:

LENGTH: 60 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: None

US-09-069-726-181

Query Match

Best Local Similarity 100.0%; Score 31; DB 14; Length 60;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

US-09-312-283A-320
Sequence 320, Application US/09312283A

GENERAL INFORMATION:

APPLICANT: Watson, James D.

APPLICANT: Strachan, Lorna

APPLICANT: Sleeman, Matthew

APPLICANT: Onrust, Rene

APPLICANT: Murison, James G.

APPLICANT: Kumble, Krishanand D.

TITLE OF INVENTION: Compositions Isolated from Skin Cells

FILE REFERENCE: 11000.1011c2

CURRENT APPLICATION NUMBER: US/09/312,283A

CURRENT FILING DATE: 1999-05-14

NUMBER OF SEQ ID NOS: 415

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 320

LENGTH: 60

TYPE: PRT

ORGANISM: Mouse

US-09-312-283A-320

Query Match

Best Local Similarity 100.0%; Score 31; DB 17; Length 60;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

US-09-312-283B-320
Sequence 320, Application US/09312283B

GENERAL INFORMATION:

APPLICANT: Watson, James D.

APPLICANT: Strachan, Lorna

APPLICANT: Sleeman, Matthew

APPLICANT: Onrust, Rene

APPLICANT: Murison, James G.

APPLICANT: Kumble, Krishanand D.

TITLE OF INVENTION: Compositions Isolated from Skin Cells

FILE REFERENCE: 11000.1011c2

CURRENT APPLICATION NUMBER: US/09/312,283B

CURRENT FILING DATE: 1999-05-14

NUMBER OF SEQ ID NOS: 425

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 181

LENGTH: 60

TYPE: PRT

ORGANISM: Mouse

US-09-312-283B-181

Query Match

Best Local Similarity 100.0%; Score 31; DB 17; Length 60;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

Y 1 GGDP 5

DB 51 GGDP 55

```

; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murlison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated from Skin Cells
; FILE REFERENCE: 11000.1011c2
; CURRENT APPLICATION NUMBER: US/09/312,283B
; NUMBER OF SEQ ID NOS: 425
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 320
; LENGTH: 60
; TYPE: PRT
; ORGANISM: Mouse
US-09-312-283B-320

```

```

Query Match
Best Local Similarity 100.0%; Score 31; DB 17; Length 60;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 GGFPD 5
DB 51 GGFPD 55

```

```

RESULT 11
US-09-312-283C-181
; Sequence 181, Application US/09312283C
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murlison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated from Skin Cells
; FILE REFERENCE: 11000.1011c2
; CURRENT APPLICATION NUMBER: US/09/312,283C
; NUMBER OF SEQ ID NOS: 425
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 181
; LENGTH: 60
; TYPE: PRT
; ORGANISM: Mouse
US-09-312-283C-181

```

```

Query Match
Best Local Similarity 100.0%; Score 31; DB 17; Length 60;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 GGFPD 5
DB 51 GGFPD 55

```

```

RESULT 12
US-09-312-283C-320
; Sequence 320, Application US/09312283C
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murlison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated from Skin Cells
; FILE REFERENCE: 11000.1011c2
; CURRENT APPLICATION NUMBER: US/09/312,283C

```

```

; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 425
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 320
; LENGTH: 60
; TYPE: PRT
; ORGANISM: Mouse
US-09-312-283C-320

```

```

Query Match
Best Local Similarity 100.0%; Score 31; DB 17; Length 60;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 GGFPD 5
DB 51 GGFPD 55

```

```

RESULT 13
US-09-866-050A-181
; Sequence 181, Application US/09866050A
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murlison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated from Skin Cells
; FILE REFERENCE: 11000.1011c4U
; CURRENT APPLICATION NUMBER: US/09/866,050A
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 181
; LENGTH: 60
; TYPE: PRT
; ORGANISM: Mouse
US-09-866-050A-181

```

```

Query Match
Best Local Similarity 100.0%; Score 31; DB 22; Length 60;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 GGFPD 5
DB 51 GGFPD 55

```

```

RESULT 14
US-09-866-050A-320
; Sequence 320, Application US/09866050A
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murlison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated from Skin Cells
; FILE REFERENCE: 11000.1011c4U
; CURRENT APPLICATION NUMBER: US/09/866,050A
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 320
; LENGTH: 60
; TYPE: PRT
; ORGANISM: Mouse
US-09-866-050A-320

```

Query Match 100.0%; Score 31; DB 22; length 60;
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

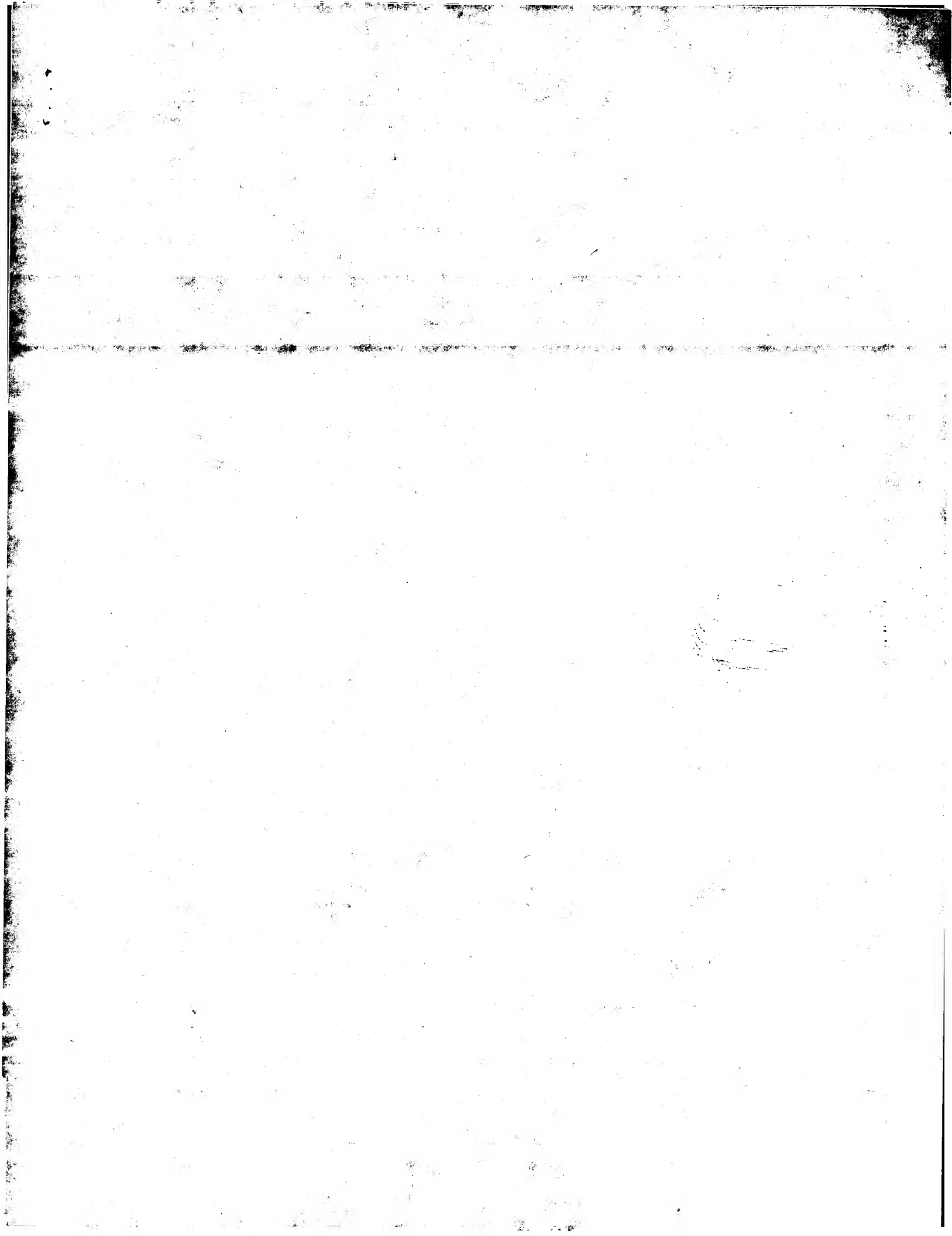
OY 1 GGDP 5
 |||||
 Db 51 GGDP 55

RESULT 15
 US-09-791-537-134706
 ; Sequence 134706, Application US/09791537
 ; GENERAL INFORMATION:
 ; APPLICANT: Bionomix, Inc.
 ; APPLICANT: Debe, Derek
 ; APPLICANT: Danzel, Joseph
 ; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
 ; FILE REFERENCE: 261/210
 ; CURRENT APPLICATION NUMBER: US/09/791,537
 ; CURRENT FILING DATE: 2001-02-22
 ; NUMBER OF SEQ ID NOS: 153055
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 134706
 ; LENGTH: 70
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-791-537-134706

Query Match 100.0%; Score 31; DB 21; length 70;
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGDP 5
 |||||
 Db 45 GGDP 49

Search completed: March 17, 2003, 12:19:35
 job time : 153 secs



GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 17, 2003, 12:09:04 ; Search time 23 Seconds
(Without alignments)
28.366 Million cell updates/sec

Title: SEQ25-FD
Perfect score: 31
Sequence: 1 gfdp 5

Scoring table: BL0SUM62
Gapop 10.0 , Gapext 0.5

Searched: 603988 seqs, 130485580 residues

Total number of hits satisfying chosen parameters: 603988

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/1/pae/PCT_US02_NEW_COMB.pep.*
2: /cgn2_6/ptodata/1/pae/US06_NEW_COMB.pep.*
3: /cgn2_6/ptodata/1/pae/US07_NEW_COMB.pep.*
4: /cgn2_6/ptodata/1/pae/US08_NEW_COMB.pep.*
5: /cgn2_6/ptodata/1/pae/US09_NEW_COMB.pep.*
6: /cgn2_6/ptodata/1/pae/US10_NEW_COMB.pep.*
7: /cgn2_6/ptodata/1/pae/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	31	100.0	11	1 PCT-US02-36496-3191	Sequence 3191, Ap
2	31	100.0	11	US-10-293-418-3191	Sequence 3191, Ap
3	31	100.0	5	US-09-855-754B-9	Sequence 9, Appl
4	31	100.0	89	1 PCT-US03-01605-293	Sequence 293, Appl
5	31	100.0	91	US-10-092-411A-5117	Sequence 5117, Ap
6	31	100.0	93	US-09-950-084-5791	Sequence 5791, Ap
7	31	100.0	98	1 PCT-US02-22011-5	Sequence 5, Appl
8	31	100.0	98	US-10-194-975-5	Sequence 5, Appl
9	31	100.0	122	1 PCT-US02-32613-33	Sequence 33, Appl
10	31	100.0	122	US-10-269-805-33	Sequence 33, Appl
11	31	100.0	131	US-09-724-676-74450	Sequence 74450, A
12	31	100.0	131	US-09-724-676-74450	Sequence 74450, A
13	31	100.0	140	US-09-724-676-74456	Sequence 74456, A
14	31	100.0	140	US-09-724-676A-74456	Sequence 74456, A
15	31	100.0	144	1 PCT-US02-32727-16148	Sequence 16148, A
16	31	100.0	144	US-09-978-825-16148	Sequence 16148, A
17	31	100.0	164	US-10-057-498-16148	Sequence 16148, A
18	31	100.0	191	US-09-724-676-74460	Sequence 74460, A
19	31	100.0	191	US-09-724-676A-74460	Sequence 74460, A
20	31	100.0	193	US-09-724-676-74454	Sequence 74454, A
21	31	100.0	193	US-09-724-676A-74454	Sequence 74454, A
22	31	100.0	197	US-10-028-248A-12	Sequence 12, Appl
23	31	100.0	197	US-10-107-782-12	Sequence 12, Appl
24	31	100.0	207	US-10-028-248A-55	Sequence 55, Appl
25	31	100.0	207	US-10-107-782-55	Sequence 55, Appl
26	31	100.0	208	1 PCT-US02-28859-88	Sequence 88, Appl

ALIGNMENTS

27	31	100.0	208	1 PCT-US02-28859-89	Sequence 89, Appl
28	31	100.0	208	US-10-241-220-88	Sequence 88, Appl
29	31	100.0	208	US-10-241-220-89	Sequence 89, Appl
30	31	100.0	208	US-10-028-248A-56	Sequence 56, Appl
31	31	100.0	208	US-10-028-248A-57	Sequence 57, Appl
32	31	100.0	208	US-10-028-248A-58	Sequence 58, Appl
33	31	100.0	208	US-10-107-782-56	Sequence 56, Appl
34	31	100.0	208	US-10-107-782-57	Sequence 57, Appl
35	31	100.0	208	US-10-107-782-58	Sequence 58, Appl
36	31	100.0	209	US-10-028-248A-10	Sequence 10, Appl
37	31	100.0	209	US-10-107-782-10	Sequence 10, Appl
38	31	100.0	248	1 PCT-US02-36496-2094	Sequence 2094, Ap
39	31	100.0	248	US-10-293-418-2094	Sequence 2094, Ap
40	31	100.0	268	US-09-724-676-74453	Sequence 74453, A
41	31	100.0	268	US-09-724-676-74459	Sequence 74459, A
42	31	100.0	268	US-09-724-676A-74453	Sequence 74453, A
43	31	100.0	268	US-09-724-676A-74459	Sequence 74459, A
44	31	100.0	268	US-10-282-122A-42841	Sequence 42841, A
45	31	100.0	301	6 US-10-366-683-33096	Sequence 33096, A

```

RESULT 1
PCT-US02-36496-3191
; Sequence 3191, Application PC/TUS0236496
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PF523PCT2
; CURRENT APPLICATION NUMBER: PCT/US02/36496
; PRIOR FILING DATE: 2002-11-14
; PRIOR APPLICATION NUMBER: 60/331,469
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 60/340,817
; PRIOR FILING DATE: 2001-12-19
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 3191
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-36496-3191

Query Match      100.0%; Score 31; DB 1; Length 11;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GGFDP 5
      |||||
Db      7 GGFDP 11

RESULT 2
US-10-293-418-3191
; Sequence 3191, Application US/10293418
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PF523P2
; CURRENT APPLICATION NUMBER: US/10/293,418
; PRIOR FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: 60/331,469
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 60/340,817
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 09/880,748
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/276,248

```

;; PRIOR FILING DATE: 2001-03-16
;; PRIOR APPLICATION NUMBER: 60/240,816
;; PRIOR FILING DATE: 2000-10-17
;; PRIOR APPLICATION NUMBER: 60/212,210
;; PRIOR FILING DATE: 2000-06-16
;; NUMBER OF SEQ ID NOS: 3247
;; SEQ ID NO 3191
;; LENGTH: 11
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-293-418-3191

Query Match
Best Local Similarity 100.0%; Score 31; DB 6; Length 11;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGFPD 5
Db 7 GGFPD 11

RESULT 3
US-09-855-754B-9
;; Sequence 9, Application US/09855754B
;; GENERAL INFORMATION:
;; APPLICANT: BOURSUX-EBUE, CAROLINE
;; TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED
;; REGIONS OF PERTACTIN IN BORDETELLA PERTUSSIS,
;; TITLE OF INVENTION: BORDETELLA PARAPERTUSSIS, AND BORDETELLA
;; TITLE OF INVENTION: BRONCHISEPTICA, THEIR USE IN DIAGNOSTICS, AND IN
;; FILE REFERENCE: 03495-0206-00000
;; CURRENT APPLICATION NUMBER: US/09/855,754B
;; PRIOR FILING DATE: 2001-05-16
;; PRIOR APPLICATION NUMBER: 60/206,969
;; PRIOR FILING DATE: 2000-05-25
;; NUMBER OF SEQ ID NOS: 25
;; SOFTWARE: Patentin Ver. 2.1
;; SEQ ID NO 9
;; LENGTH: 56
;; TYPE: PRT
;; ORGANISM: Bordetella bronchiseptica
US-09-855-754B-9

Query Match
Best Local Similarity 100.0%; Score 31; DB 5; Length 56;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGFPD 5
Db 23 GGFPD 27

RESULT 4
PCT-US03-01605-293
;; Sequence 293, Application PC/TUS0301605
;; GENERAL INFORMATION:
;; APPLICANT: INCYTE GENOMICS, INC.
;; APPLICANT: JONES, Anissa L.; DAHL, Christopher R.;
;; APPLICANT: GIETZEN, Darryl; CHANN, Joyce;
;; APPLICANT: DUFOR, Gerard E.; JACKSON, Jennifer L.;
;; APPLICANT: YU, Jimmy Y.; TIVANSON, Olivia;
;; APPLICANT: YAP, Pierre E.; AMSHEY, Stefan R.;
;; APPLICANT: DAM, Tam C.; LIU, Tommy F.;
;; APPLICANT: GERSTIN JR., Edward H.; PERALTA, Careyna H.;
;; APPLICANT: LEWIS, Samantha A.; CHEN, Alice;
;; APPLICANT: MARWAHA, Rakesh; LAN, Ruth Y.;
;; APPLICANT: URASHKA, Michael E.; KRISTYAN, Steenivasa R.;
;; APPLICANT: KOLJURO, Vijaykumar; PANESAR, Iqbal S.
;; TITLE OF INVENTION: SECRETORY MOLECULES
;; FILE REFERENCE: PT-1241 PCT
;; CURRENT APPLICATION NUMBER: PCT/US03/01605

;; CURRENT FILING DATE: 2003-01-15
;; PRIOR APPLICATION NUMBER: US 60/349,946
;; PRIOR FILING DATE: 2002-01-17
;; PRIOR APPLICATION NUMBER: US 60/349,413
;; PRIOR FILING DATE: 2002-01-17
;; NUMBER OF SEQ ID NOS: 324
;; SOFTWARE: PERL Program
;; SEQ ID NO 293
;; LENGTH: 89
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc_feature
;; OTHER INFORMATION: Incyte ID No: IG:474421.5.orf3:2002JANI8
PCT-US03-01605-293

Query Match
Best Local Similarity 100.0%; Score 31; DB 1; Length 89;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGFPD 5
Db 61 GGFPD 65

RESULT 5
US-10-092-411A-5117
;; Sequence 5117, Application US/10092411A
;; GENERAL INFORMATION:
;; APPLICANT: Lynn Doucette-Stamm et al
;; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
;; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
;; FILE REFERENCE: 032796-101
;; CURRENT APPLICATION NUMBER: US/10/092,411A
;; PRIOR FILING DATE: 2002-03-07
;; PRIOR APPLICATION NUMBER: US 09/134,001
;; PRIOR FILING DATE: 1998-08-13
;; PRIOR APPLICATION NUMBER: US 60/064,964
;; PRIOR FILING DATE: 1997-11-08
;; PRIOR APPLICATION NUMBER: US 60/055,779
;; PRIOR FILING DATE: 1997-08-14
;; NUMBER OF SEQ ID NOS: 5676
;; SEQ ID NO 5117
;; LENGTH: 91
;; TYPE: PRT
;; ORGANISM: Staphylococcus epidermidis
US-10-092-411A-5117

Query Match
Best Local Similarity 100.0%; Score 31; DB 6; Length 91;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGFPD 5
Db 72 GGFPD 76

RESULT 6
US-09-950-084-5791
;; Sequence 5791, Application US/09950084
;; GENERAL INFORMATION:
;; APPLICANT: George H. Shimer, Jr.
;; APPLICANT: George H. Miller
;; APPLICANT: Robert S. Hare
;; APPLICANT: Karen J. Shaw
;; TITLE OF INVENTION: Staphylococcus aureus Related Compositions and Methods
;; FILE REFERENCE: 1034/1C8630S2
;; CURRENT APPLICATION NUMBER: US/09/950,084
;; PRIOR FILING DATE: 2001-09-10
;; PRIOR APPLICATION NUMBER: US 09/417,811
;; PRIOR FILING DATE: 1999-10-14
;; PRIOR APPLICATION NUMBER: US 09/353,718
;; PRIOR FILING DATE: 1999-07-14

```
; PRIOR APPLICATION NUMBER: US 09/266,557
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,556
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,555
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,542
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,541
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/037,934
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: US 09/036,720
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/036,338
; PRIOR FILING DATE: 1998-03-06
; Remaining Prior Application data removed - See File Wrapper or PALM.
; SEQ ID NO 5791
; LENGTH: 93
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-950-084-5791
```

```
Query Match          100.0%; Score 31; DB 5; Length 93;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 GGFPD 5
      |||||
Db 71 GGFPD 75
```

```
RESULT 7
PCT-US02-22011-5
; Sequence 5, Application PC/TUS0222011
; GENERAL INFORMATION:
; APPLICANT: Foote, Jefferson
; TITLE OF INVENTION: Super Humanized Antibodies
; FILE REFERENCE: 501231.02
; CURRENT APPLICATION NUMBER: PCT/US02/22011
; CURRENT FILING DATE: 2002-07-12
; PRIOR APPLICATION NUMBER: US 60/305,111
; PRIOR FILING DATE: 2001-07-12
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-22011-5
```

```
Query Match          100.0%; Score 31; DB 1; Length 98;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 GGFPD 5
      |||||
Db 49 GGFPD 53
```

```
RESULT 8
US-10-194-975-5
; Sequence 5, Application US/10194975
; GENERAL INFORMATION:
; APPLICANT: Foote, Jefferson
; TITLE OF INVENTION: Super Humanized Antibodies
; FILE REFERENCE: 501231.01
; CURRENT APPLICATION NUMBER: US/10/194,975
; CURRENT FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: US 60/305,111
; PRIOR FILING DATE: 2001-07-12
; NUMBER OF SEQ ID NOS: 122
```

```
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-194-975-5
```

```
Query Match          100.0%; Score 31; DB 6; Length 98;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 GGFPD 5
      |||||
Db 49 GGFPD 53
```

```
RESULT 9
PCT-US02-32613-33
; Sequence 33, Application PC/TUS0232613
; GENERAL INFORMATION:
; APPLICANT: AMGEN INC.
; TITLE OF INVENTION: ANGIOPOIETIN-2 SPECIFIC BINDING AGENTS
; FILE REFERENCE: A-722
; CURRENT APPLICATION NUMBER: PCT/US02/32613
; CURRENT FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: US 60/328,604
; PRIOR FILING DATE: 2001-10-11
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 33
; LENGTH: 122
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-32613-33
```

```
Query Match          100.0%; Score 31; DB 1; Length 122;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 GGFPD 5
      |||||
Db 49 GGFPD 53
```

```
RESULT 10
US-10-269-805-33
; Sequence 33, Application US/10269805
; GENERAL INFORMATION:
; APPLICANT: OLINER, JONATHAN D.
; TITLE OF INVENTION: ANGIOPOIETIN-2 SPECIFIC BINDING AGENTS
; FILE REFERENCE: A-722
; CURRENT APPLICATION NUMBER: US/10/269,805
; CURRENT FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: US 60/328,604
; PRIOR FILING DATE: 2001-10-11
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 33
; LENGTH: 122
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-269-805-33
```

```
Query Match          100.0%; Score 31; DB 6; Length 122;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 GGFPD 5
      |||||
Db 49 GGFPD 53
```

```
RESULT 11
```

US-09-724-676-74450
; Sequence 74450, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 74450
; LENGTH: 131
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-74450

Query Match 100.0%; Score 31; DB 5; Length 131;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFPD 5
Db 57 GGFPD 61

RESULT 12
US-09-724-676A-74450
; Sequence 74450, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 74450
; LENGTH: 131
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-74450

Query Match 100.0%; Score 31; DB 5; Length 131;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFPD 5
Db 57 GGFPD 61

RESULT 13
US-09-724-676-74456
; Sequence 74456, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 74456
; LENGTH: 140
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-74456

Query Match 100.0%; Score 31; DB 5; Length 140;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFPD 5

Db 57 GGFPD 61

RESULT 14
US-09-724-676A-74456
; Sequence 74456, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 74456
; LENGTH: 140
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-74456

Query Match 100.0%; Score 31; DB 5; Length 140;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFPD 5
Db 57 GGFPD 61

RESULT 15
PCT-US02-32727-16148
; Sequence 16148, Application PC/TUS0232727
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Skeiky, Yasir
; APPLICANT: Persing, David
; APPLICANT: Bhatia, Ajay
; APPLICANT: Maisonneuve, Jean Francois
; APPLICANT: Zhang, Yanni
; APPLICANT: Wang, Siqing
; APPLICANT: Jen, Shylan
; APPLICANT: Lodes, Michael
; APPLICANT: Benson, Darin
; APPLICANT: Jones, Robert
; APPLICANT: Carter, Derrick
; APPLICANT: Barth, Brenda
; APPLICANT: Douglass, John
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne
; FILE REFERENCE: 210121.514C1
; CURRENT APPLICATION NUMBER: PCT/US02/32727
; CURRENT FILING DATE: 2002-10-11
; NUMBER OF SEQ ID NOS: 30992
; SEQ ID NO 16148
; LENGTH: 164
; TYPE: PRT
; ORGANISM: Propionl acnes
PCT-US02-32727-16148

Query Match 100.0%; Score 31; DB 1; Length 164;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFPD 5
Db 61 GGFPD 65

Search completed: March 17, 2003, 12:20:52
Job time : 24 secs

GenCore version 5.1.4.p5-4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 17, 2003, 12:05:19 ; Search time 31.3333 Seconds
(Without alignments)
21.263 Million cell updates/sec

Title: SEQ25-FG
Perfect score: 31.3333
Sequence: 115 995 995

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002.*

1:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31	100.0	10	23	AAE17142
2	31	100.0	15	23	AAE17145
3	31	100.0	15	23	AAE17163
4	31	100.0	15	23	AAE17165
5	31	100.0	15	23	AAE17166
6	31	100.0	15	23	AAE17167
7	31	100.0	20	23	AAE17144
8	31	100.0	26	22	AB550677
9	31	100.0	31	23	AAE17173
10	31	100.0	36	23	AAE17141

11	31	100.0	36	23	AAE17170
12	31	100.0	36	23	AAE17172
13	31	100.0	36	23	AAE17174
14	31	100.0	41	23	AAE17171
15	31	100.0	45	23	AAE17153
16	31	100.0	45	23	AAE17154
17	31	100.0	45	23	AAE17155
18	31	100.0	45	23	AAE17156
19	31	100.0	45	23	AAE17157
20	31	100.0	45	23	AAE17158
21	31	100.0	45	23	AAE17159
22	31	100.0	51	23	AAE16187
23	31	100.0	51	23	AAE16186
24	31	100.0	54	23	AAE16192
25	31	100.0	54	23	AAE17150
26	31	100.0	54	23	AAE17151
27	31	100.0	56	23	AAE16188
28	31	100.0	56	23	AAE16189
29	31	100.0	56	23	AAE16191
30	31	100.0	56	23	AAE17152
31	31	100.0	59	23	AAE17147
32	31	100.0	59	23	AAE17149
33	31	100.0	61	23	AAE16190
34	31	100.0	64	23	AAE17148
35	31	100.0	69	23	ABP3149
36	31	100.0	74	22	AAU45938
37	31	100.0	80	22	AAU42855
38	31	100.0	117	14	AAK38612
39	31	100.0	122	23	ABP34634
40	31	100.0	124	22	AAO05161
41	31	100.0	126	22	ABG16939
42	31	100.0	132	22	AAO03167
43	31	100.0	135	22	AAU58992
44	31	100.0	137	22	AAU92718
45	31	100.0	140	22	AAU87608

ALIGNMENTS

RESULT 1
ID AAE17142 standard; peptide; 10 AA.
XX AAE17142:
AC
AC AAE17142:
XX
XX 18-APR-2002 (first entry)
XX
XX Bordetella pertussis pertactin (Prn1) related peptide #1.
DE
DE Pertactin: prn1; antibacterial; immunostimulant; vaccine;
XX
XX diptheria; tetanus; polio; Haemophilus influenza b infection; therapy;
KW
KW Immune response.
XX
XX Unidentified.
OS
OS WO200200695-A2.
XX
XX
XX
XX 03-JAN-2002.
PD
PD 29-JUN-2001; 2001WO-NL00493.
XX
XX 30-JUN-2000; 2000EP-0202309.
PR
PR (NME-) NEDERLANDEN MIN WELZON.
PA
PA WO01 FR:
XX
XX WO01 FR:
XX
XX WPI; 2002-139897/18.
DR
DR
XX
XX New polypeptides derived from Bordetella pertussis pertactin, useful as
PT a vaccine against infections caused by Bordetella strains, and other
PT Infectious diseases of mammals, e.g. diptheria, tetanus, or polio

XX PS Disclosure; Page 7; 52pp; English.
 CC CC The invention relates to polypeptides derived from Bordetella pertussis
 CC CC pertactin (Prn1). The polypeptide is useful in the preparation of
 CC CC vaccines against B. pertussis, B. parapertussis, B. bronchiseptica and
 CC CC other infectious diseases of mammals including diphtheria, tetanus, polio
 CC CC and infections caused by Haemophilus influenza b. The polypeptide is
 CC CC especially useful for eliciting an immune response against Bordetella sp.
 CC CC Antibodies against the polypeptide may be used for pharmaceutical and/or
 CC CC diagnostic purposes, particularly for treating or preventing infections
 CC CC caused by Bordetella pertussis or Bordetella parapertussis. The present
 CC CC sequence is B. pertussis prn1 related peptide used in the invention.
 SQ Sequence 10 AA;
 QY Query Match 100.0%; Score 31; DB 23; Length 10;
 DB Best Local Similarity 100.0%; Pred. No. 17;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GFGFP 5
 DB 1 GFGFP 5
 RESULT 2
 ID AAE17145 standard; peptide; 15 AA.
 AC AAE17145;
 XX
 DT 18-APR-2002 (first entry)
 XX
 DE Bordetella pertussis pertactin (Prn1) related peptide #4.
 KW Pertactin; prn1; antibacterial; immunostimulant; antimicrobial; vaccine;
 KW diphtheria; tetanus; polio; Haemophilus influenza b infection; therapy;
 KW immune response.
 XX
 OS Unidentified.
 XX
 PN WO200200695-A2.
 PD 03-JAN-2002.
 XX
 PF 29-JUN-2001; 2001WO-NL00493.
 PR 30-JUN-2000; 2000EP-0202309.
 XX
 PA (NENE-) NEDERLANDEN MIN WELZIJN.
 PI
 MOOI FR:
 WPI; 2002-139897/18.
 DR
 XX
 PT New polypeptides derived from Bordetella pertussis pertactin, useful as
 PT a vaccine against infections caused by Bordetella strains, and other
 PT infectious diseases of mammals, e.g. diphtheria, tetanus, or polio
 PS
 SQ Claim 6; Page 7; 52pp; English.
 CC The invention relates to polypeptides derived from Bordetella pertussis
 CC pertactin (Prn1). The polypeptide is useful in the preparation of
 CC vaccines against B. pertussis, B. parapertussis, B. bronchiseptica and
 CC other infectious diseases of mammals including diphtheria, tetanus, polio
 CC and infections caused by Haemophilus influenza b. The polypeptide is
 CC especially useful for eliciting an immune response against Bordetella sp.
 CC Antibodies against the polypeptide may be used for pharmaceutical and/or
 CC diagnostic purposes, particularly for treating or preventing infections
 CC caused by Bordetella pertussis or Bordetella parapertussis. The present
 CC sequence is B. pertussis prn1 related peptide used in the invention.
 SQ Sequence 15 AA;

Query Match 100.0%; Score 31; DB 23; Length 15;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GFGFP 5
 DB 6 GFGFP 10
 RESULT 3
 ID AAE17163 standard; peptide; 15 AA.
 AC AAE17163;
 XX
 DT 18-APR-2002 (first entry)
 XX
 DE Monoclonal Ab derived pep4 peptide specific to region 1 of pertactin.
 XX
 KW Pertactin; prn1; antibacterial; immunostimulant; antimicrobial; vaccine;
 KW diphtheria; tetanus; polio; Haemophilus influenza b infection; therapy;
 KW immune response.
 XX
 OS Unidentified.
 XX
 PN WO200200695-A2.
 PD 03-JAN-2002.
 XX
 PF 29-JUN-2001; 2001WO-NL00493.
 PR 30-JUN-2000; 2000EP-0202309.
 XX
 PA (NENE-) NEDERLANDEN MIN WELZIJN.
 PI
 MOOI FR:
 WPI; 2002-139897/18.
 DR
 XX
 PT New polypeptides derived from Bordetella pertussis pertactin, useful as
 PT a vaccine against infections caused by Bordetella strains, and other
 PT infectious diseases of mammals, e.g. diphtheria, tetanus, or polio
 PS
 SQ Example 1; Fig 2b; 52pp; English.
 CC The invention relates to polypeptides derived from Bordetella pertussis
 CC pertactin (Prn1). The polypeptide is useful in the preparation of
 CC vaccines against B. pertussis, B. parapertussis, B. bronchiseptica and
 CC other infectious diseases of mammals including diphtheria, tetanus, polio
 CC and infections caused by Haemophilus influenza b. The polypeptide is
 CC especially useful for eliciting an immune response against Bordetella sp.
 CC Antibodies against the polypeptide may be used for pharmaceutical and/or
 CC diagnostic purposes, particularly for treating or preventing infections
 CC caused by Bordetella pertussis or Bordetella parapertussis. The present
 CC sequence is a peptide derived from monoclonal antibody which is specific
 CC to region 1 of pertactin.
 SQ Sequence 15 AA;
 QY Query Match 100.0%; Score 31; DB 23; Length 15;
 DB Best Local Similarity 100.0%; Pred. No. 24;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GFGFP 5
 DB 11 GFGFP 15
 RESULT 4
 ID AAE17165 standard; peptide; 15 AA.

```

AC  AAE17165;
XX
XX  18-APR-2002 (first entry)
DE  Monoclonal Ab derived Pep6 peptide specific to region 1 of pertactin.
XX
XX  Pertactin; prn1; antibacterial; immunostimulant; antimicrobial; vaccine;
KM  diphtheria; tetanus; polio; Haemophilus influenza b infection; therapy;
XX  immune response.
XX  Unidentified.
XX  OS
XX  PN  WO200200695-A2.
XX
XX  03-JAN-2002.
XX
XX  29-JUN-2001; 2001WO-NL00493.
XX
XX  30-JUN-2000; 2000EP-0202309.
XX
XX  (NEME-) NEDERLANDEN MIN WELZIJN.
XX
XX  Mool FR;
XX
XX  WPI; 2002-139897/18.
XX
XX  New polypeptides derived from Bordetella pertussis pertactin, useful as
PT  a vaccine against infections caused by Bordetella strains, and other
PT  infectious diseases of mammals, e.g. diphtheria, tetanus, or polio
XX
XX  Example 1; Fig 2b; 52pp; English.
XX
XX  The invention relates to polypeptides derived from Bordetella pertussis
CC  pertactin (Prn1). The polypeptide is useful in the preparation of
CC  vaccines against B. pertussis, B. parapertussis, B. bronchiseptica and
CC  other infectious diseases of mammals including diphtheria, tetanus, polio
CC  and infections caused by Haemophilus influenza b. The polypeptide is
CC  especially useful for eliciting an immune response against Bordetella sp.
CC  Antibodies against the polypeptide may be used for pharmaceutical and/or
CC  diagnostic purposes, particularly for treating or preventing infections
CC  caused by Bordetella pertussis or Bordetella parapertussis. The present
CC  sequence is a peptide derived from monoclonal antibody which is specific
CC  to region 1 of pertactin.
XX
XX  Sequence 15 AA;
SQ
XX
XX  Query Match 100.0%; Score 31; DB 23; Length 15;
XX  Best Local Similarity 100.0%; Pred. No. 24;
XX  Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY  1 GGRGP 5
XX  |||||
DB  1 GGRGP 5
XX
XX  RESULT 5
XX  AAE17166
XX  ID  AAE17166 standard; peptide; 15 AA.
XX
XX  AAE17166;
XX
XX  18-APR-2002 (first entry)
XX
XX  Monoclonal Ab derived Pep7 peptide specific to region 1 of pertactin.
DE  Pertactin; prn1; antibacterial; immunostimulant; antimicrobial; vaccine;
KM  diphtheria; tetanus; polio; Haemophilus influenza b infection; therapy;
XX  immune response.
XX  Unidentified.
XX  OS
XX  PN  WO200200695-A2.
XX
XX

```

```

PD  03-JAN-2002.
XX
XX  29-JUN-2001; 2001WO-NL00493.
XX
XX  30-JUN-2000; 2000EP-0202309.
XX
XX  (NEME-) NEDERLANDEN MIN WELZIJN.
XX
XX  Mool FR;
XX
XX  WPI; 2002-139897/18.
XX
XX  New polypeptides derived from Bordetella pertussis pertactin, useful as
PT  a vaccine against infections caused by Bordetella strains, and other
PT  infectious diseases of mammals, e.g. diphtheria, tetanus, or polio
XX
XX  Example 1; Fig 2b; 52pp; English.
XX
XX  The invention relates to polypeptides derived from Bordetella pertussis
CC  pertactin (Prn1). The polypeptide is useful in the preparation of
CC  vaccines against B. pertussis, B. parapertussis, B. bronchiseptica and
CC  other infectious diseases of mammals including diphtheria, tetanus, polio
CC  and infections caused by Haemophilus influenza b. The polypeptide is
CC  especially useful for eliciting an immune response against Bordetella sp.
CC  Antibodies against the polypeptide may be used for pharmaceutical and/or
CC  diagnostic purposes, particularly for treating or preventing infections
CC  caused by Bordetella pertussis or Bordetella parapertussis. The present
CC  sequence is a peptide derived from monoclonal antibody which is specific
CC  to region 1 of pertactin.
XX
XX  Sequence 15 AA;
SQ
XX
XX  Query Match 100.0%; Score 31; DB 23; Length 15;
XX  Best Local Similarity 100.0%; Pred. No. 24;
XX  Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY  1 GGRGP 5
XX  |||||
DB  1 GGRGP 5
XX
XX  RESULT 6
XX  AAE17167
XX  ID  AAE17167 standard; peptide; 15 AA.
XX
XX  AAE17167;
XX
XX  18-APR-2002 (first entry)
XX
XX  Monoclonal Ab derived Pep10 peptide specific to region 1 of pertactin.
DE  Pertactin; prn1; antibacterial; immunostimulant; antimicrobial; vaccine;
KM  diphtheria; tetanus; polio; Haemophilus influenza b infection; therapy;
XX  immune response.
XX  Unidentified.
XX  OS
XX  PN  WO200200695-A2.
XX
XX  03-JAN-2002.
XX
XX  29-JUN-2001; 2001WO-NL00493.
XX
XX  30-JUN-2000; 2000EP-0202309.
XX
XX  (NEME-) NEDERLANDEN MIN WELZIJN.
XX
XX  Mool FR;
XX
XX  WPI; 2002-139897/18.
XX
XX  New polypeptides derived from Bordetella pertussis pertactin, useful as
PT  a vaccine against infections caused by Bordetella strains, and other
PT

```

PT Infectious diseases of mammals, e.g. diphtheria, tetanus, or polio
 XX
 PS Example 1; Fig 2b; 52pp; English.
 CC The invention relates to polypeptides derived from *Bordetella pertussis*
 CC pertactin (Prn1). The polypeptide is useful in the preparation of
 CC vaccines against *B. pertussis*, *B. parapertussis*, *B. bronchiseptica* and
 CC other infectious diseases of mammals including diphtheria, tetanus, polio
 CC and infections caused by *Haemophilus influenza* b. The polypeptide is
 CC especially useful for eliciting an immune response against *Bordetella* sp.
 CC Antibodies against the polypeptide may be used for pharmaceutical and/or
 CC diagnostic purposes, particularly for treating or preventing infections
 CC caused by *Bordetella pertussis* or *Bordetella parapertussis*. The present
 CC sequence is a peptide derived from monoclonal antibody which is specific
 CC to region 1 of pertactin.
 SQ Sequence 15 AA;
 Query Match 100.0%; Score 31; DB 23; Length 15;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGRGP 5
 DB 1 GGRGP 5
 RESULT 7
 AAEL17144
 ID AAEL17144 standard; peptide; 20 AA.
 XX
 AC AAEL17144;
 XX
 DT 18-APR-2002 (first entry)
 XX
 DE *Bordetella pertussis* pertactin (Prn1) related peptide #3.
 XX
 KW Pertactin; prn1; antibacterial; immunostimulant; antimicrobial; vaccine;
 KW diphtheria; tetanus; polio; *Haemophilus influenza* b infection; therapy;
 KW immune response.
 XX
 OS Unidentified.
 XX
 PN WO200200695-A2.
 XX
 PD 03-JAN-2002.
 XX
 PF 29-JUN-2001; 2001WO-NL00493.
 XX
 PR 30-JUN-2000; 2000EP-0202309.
 XX
 PA (NEW-) NEDERLANDEN MIN WELZIJN.
 XX
 PI MOOI FR;
 XX
 DR WPI; 2002-139897/18.
 XX
 PT New polypeptides derived from *Bordetella pertussis* pertactin, useful as
 PT a vaccine against infections caused by *Bordetella* strains, and other
 PT infectious diseases of mammals, e.g. diphtheria, tetanus, or polio
 XX
 PS Disclosure; Page 7; 52pp; English.
 XX
 CC The invention relates to polypeptides derived from *Bordetella pertussis*
 CC pertactin (Prn1). The polypeptide is useful in the preparation of
 CC vaccines against *B. pertussis*, *B. parapertussis*, *B. bronchiseptica* and
 CC other infectious diseases of mammals including diphtheria, tetanus, polio
 CC and infections caused by *Haemophilus influenza* b. The polypeptide is
 CC especially useful for eliciting an immune response against *Bordetella* sp.
 CC Antibodies against the polypeptide may be used for pharmaceutical and/or
 CC diagnostic purposes, particularly for treating or preventing infections
 CC caused by *Bordetella pertussis* or *Bordetella parapertussis*. The present
 CC sequence is *B. pertussis* prn1 related peptide used in the invention.

XX
 SQ Sequence 20 AA;
 Query Match 100.0%; Score 31; DB 23; Length 20;
 Best Local Similarity 100.0%; Pred. No. 31;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGRGP 5
 DB 11 GGRGP 15
 RESULT 8
 ABB50677
 ID ABB50677 standard; Protein; 26 AA.
 XX
 AC ABB50677;
 XX
 DT 07-FEB-2002 (first entry)
 XX
 DE Human secreted protein encoded by gene 45 SEQ ID NO:625.
 XX
 KW Human; secreted protein; immunomodulatory; antisclerotic; anti-HIV;
 KW dermatological; immunosuppressive; antiinflammatory; immunostimulant;
 KW cytostatic; cardiac; vascular; anti-angiogenic; ophthalmological;
 KW neuroprotective; neurotropic; anticonvulsant; antialzheimers; vulnery;
 KW antiparkinsonian; antimicrobial; gene therapy; vaccine; immune disorder;
 KW multiple sclerosis; systemic lupus erythematosus; HIV infection; cancer;
 KW Gaucher's disease; cardiovascular disease; Scleritar syndrome; chemotaxis;
 KW Chaga's cardiomyopathy; coronary arteriosclerosis; angogenic disorder;
 KW corneal graft neovascularisation; diabetes; retinopathy; regeneration;
 KW neurological disorder; Huntington's chorea; Alzheimer's disease;
 KW Parkinson's disease; infectious disease.
 XX
 OS Homo sapiens.
 XX
 PN WO200162891-A2.
 XX
 PD 30-AUG-2001.
 XX
 PF 21-FEB-2001; 2001WO-US05614.
 XX
 PR 24-FEB-2000; 2000US-184836P.
 XX
 PR 29-MAR-2000; 2000US-193170P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Ni J, Edner R, Lafleur DW, Moore PA, Olsen HS, Rosen CA;
 PI Ruben SM, Soppet DR, Young PE, Shi Y, Florence KA, Wei Y;
 PI Florence C, Hu J, Li Y, Kyaw H, Fischer CL, Ferris AM, Fan P;
 PI Feng P, Endress GA, Dillon PJ, Carter KC, Brewer LA, Yu G;
 PI Zeng Z, Greene JM;
 XX
 DR WPI; 2001-625724/72.
 XX
 PT Nucleic acids encoding 207 human secreted polypeptides, useful for
 PT preventing, diagnosing and/or treating, e.g. cancers, Parkinson's
 PT disease and diabetic retinopathy
 XX
 PS Disclosure; Page 105; 1533pp; English.
 XX
 CC ABB50301 to ABB51287 and ABA83194 to ABA83441 represent human secreted
 CC proteins (I) and polynucleotide (II) sequences. (I) and (II) have various
 CC activities based on the tissues and cells the genes are expressed in.
 CC Example of these activities include: immunomodulatory; antisclerotic;
 CC dermatological; immunosuppressive; antiinflammatory; immunostimulant;
 CC anti-HIV; cytostatic; cardiac; anti-angiogenic; ophthalmological;
 CC neuroprotective; neurotropic; anticonvulsant; antialzheimers; vascular;
 CC antiparkinsonian; antimicrobial; gene therapy; vaccine; immune disorder;
 CC multiple sclerosis; systemic lupus erythematosus; HIV infection; cancer;
 CC Gaucher's disease; cardiovascular disease; Scleritar syndrome; chemotaxis;
 CC Chaga's cardiomyopathy; coronary arteriosclerosis; angogenic disorder;
 CC corneal graft neovascularisation; diabetes; retinopathy; regeneration;
 CC neurological disorder; Huntington's chorea; Alzheimer's disease;
 CC Parkinson's disease; infectious disease.
 CC

CC (HIV) infections), hyperproliferative disorders (e.g. cancers and
 CC Gaucher's disease), cardiovascular diseases (e.g. Scleritar syndrome,
 CC Chaga's cardiomyopathy and coronary arteriosclerosis), angiotonic
 CC disorders (e.g. corneal graft neovascularisation and diabetic
 CC retinopathy), neurological disorders (e.g. Huntington's chorea,
 CC Alzheimer's disease and Parkinson's disease), infectious diseases and/or
 CC for promoting wound healing, regeneration and/or chemotaxis. ABA83185 to
 CC ABA83193 and ABA850300 represent sequences used in the exemplification of
 CC the present invention.

XX Sequence 26 AA;

Query Match 100.0%; Score 31; DB 22; Length 26;

Best Local Similarity 100.0%; Pred. No. 40;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFCP 5
 |||||
 DB 18 GGFCP 22

RESULT 9

ID AAE17173 standard; Protein; 31 AA.

AC AAE17173;

DT 18-APR-2002 (first entry)

DE Maltose binding protein-pertactin region 1 fusion protein, MBP-Prt4.

KW Pertactin; prn1; antibacterial; immunostimulant; antimicrobial; vaccine;

KM diphtheria; tetanus; polio; Haemophilus influenza b infection; therapy;

XX Immune response; maltose binding protein; MBP; fusion protein.

OS Chimeric - Bordetella pertussis.

PN WO200200695-A2.

PD 03-JAN-2002.

PF 29-JUN-2001; 2001MO-NL00493.

PR 30-JUN-2000; 2000EP-0202309.

PA (NEME-) NEDERLANDEN MIN WELZIJN.

PI Mool FR;

DR WPI; 2002-139897/18.

XX New polypeptides derived from Bordetella pertussis pertactin, useful as

PT a vaccine against infections caused by Bordetella strains, and other

CC infectious diseases of mammals, e.g. diphtheria, tetanus, or polio

XX Example 1; Fig 4b; 52pp; English.

CC The invention relates to polypeptides derived from Bordetella pertussis

CC pertactin (Prn1). The polypeptide is useful in the preparation of

CC vaccines against B. pertussis, B. parapertussis, B. bronchiseptica and

CC other infectious diseases of mammals including diphtheria, tetanus, polio

CC and infections caused by Haemophilus influenza b. The polypeptide is

CC especially useful for eliciting an immune response against Bordetella sp.

CC Antibodies against the polypeptide may be used for pharmaceutical and/or

CC diagnostic purposes, particularly for treating or preventing infections

CC caused by Bordetella pertussis or Bordetella parapertussis. The present

CC sequence is maltose binding protein (MBP)-pertactin region 1 fusion

CC protein used in the exemplification of the invention.

XX Sequence 31 AA;

XX Query Match 100.0%; Score 31; DB 23; Length 31;

Best Local Similarity 100.0%; Pred. No. 46;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFCP 5
 |||||
 DB 20 GGFCP 24

RESULT 10

ID AAE17141 standard; peptide; 36 AA.

AC AAE17141;

DT 18-APR-2002 (first entry)

DE Bordetella pertussis pertactin (Prn1) peptide.

KW Pertactin; prn1; antibacterial; immunostimulant; antimicrobial; vaccine;

KM diphtheria; tetanus; polio; Haemophilus influenza b infection; therapy;

XX Immune response.

OS Bordetella pertussis.

PN WO200200695-A2.

PD 03-JAN-2002.

PF 29-JUN-2001; 2001MO-NL00493.

PR 30-JUN-2000; 2000EP-0202309.

PA (NEME-) NEDERLANDEN MIN WELZIJN.

PI Mool FR;

DR WPI; 2002-139897/18.

XX New polypeptides derived from Bordetella pertussis pertactin, useful as

PT a vaccine against infections caused by Bordetella strains, and other

CC infectious diseases of mammals, e.g. diphtheria, tetanus, or polio

XX Claim 3; Page 4; 52pp; English.

CC The invention relates to polypeptides derived from Bordetella pertussis

CC pertactin (Prn1). The polypeptide is useful in the preparation of

CC vaccines against B. pertussis, B. parapertussis, B. bronchiseptica and

CC other infectious diseases of mammals including diphtheria, tetanus, polio

CC and infections caused by Haemophilus influenza b. The polypeptide is

CC especially useful for eliciting an immune response against Bordetella sp.

CC Antibodies against the polypeptide may be used for pharmaceutical and/or

CC diagnostic purposes, particularly for treating or preventing infections

CC caused by Bordetella pertussis or Bordetella parapertussis. The present

XX sequence is B. pertussis prn1 peptide.

XX Example 1; Fig 4b; 52pp; English.

CC The invention relates to polypeptides derived from Bordetella pertussis

CC pertactin (Prn1). The polypeptide is useful in the preparation of

CC vaccines against B. pertussis, B. parapertussis, B. bronchiseptica and

CC other infectious diseases of mammals including diphtheria, tetanus, polio

CC and infections caused by Haemophilus influenza b. The polypeptide is

CC especially useful for eliciting an immune response against Bordetella sp.

CC Antibodies against the polypeptide may be used for pharmaceutical and/or

CC diagnostic purposes, particularly for treating or preventing infections

CC caused by Bordetella pertussis or Bordetella parapertussis. The present

CC sequence is B. pertussis prn1 peptide.

XX Example 1; Fig 4b; 52pp; English.

XX Query Match 100.0%; Score 31; DB 23; Length 36;

XX Best Local Similarity 100.0%; Pred. No. 53;

XX Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFCP 5
 |||||
 DB 25 GGFCP 29

RESULT 11

ID AAE17170 standard; Protein; 36 AA.

AC AAE17170;

DT 18-APR-2002 (first entry)

XX DE Maltose binding protein-pertactin region 1 fusion protein, MBP-Prt1.
 XX DE Pertactin; prt1; antibacterial; immunostimulant; antimicrobial; vaccine;
 KW diphtheria; tetanus; polio; Haemophilus influenza b infection; therapy;
 KW immune response; maltose binding protein; MBP; fusion protein.
 XX OS Chimeric - Bordetella pertussis.
 OS Chimeric - Unidentified.
 XX PN WO200200695-A2.
 XX PD 03-JAN-2002.
 XX PF 29-JUN-2001; 2001WO-NL00493.
 XX PR 30-JUN-2000; 2000EP-0202309.
 XX PA (NEME-) NEDERLANDEN MIN WELZIJN.
 XX PI Mooi FR;
 XX DR WPI; 2002-139897/18.
 XX PT New polypeptides derived from Bordetella pertussis pertactin, useful as
 PT a vaccine against infections caused by Bordetella strains, and other
 PT infectious diseases of mammals, e.g. diphtheria, tetanus, or polio -
 XX PS Example 1; Fig 4b; 52pp; English.
 XX CC The invention relates to polypeptides derived from Bordetella pertussis
 CC pertactin (Prt1). The polypeptide is useful in the preparation of
 CC vaccines against B. pertussis, B. parapertussis, B. bronchiseptica and
 CC other infectious diseases of mammals including diphtheria, tetanus, polio
 CC and infections caused by Haemophilus influenza b. The polypeptide is
 CC especially useful for eliciting an immune response against Bordetella sp.
 CC Antibodies against the polypeptide may be used for pharmaceutical and/or
 CC diagnostic purposes, particularly for treating or preventing infections
 CC caused by Bordetella pertussis or Bordetella parapertussis. The present
 CC sequence is maltose binding protein (MBP)-pertactin region 1 fusion
 CC protein used in the exemplification of the invention.
 XX SQ Sequence 36 AA;
 Query Match 100.0%; Score 31; DB 23; Length 36;
 Best Local Similarity 100.0%; Pred. No. 53;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGFPG 5
 DB 25 GGFPG 29
 DE 18-APR-2002 (first entry)
 AC AAE17172;
 AC AAE17172;
 DE Maltose binding protein-pertactin region 1 fusion protein, MBP-Prt3.
 DE Pertactin; prt1; antibacterial; immunostimulant; antimicrobial; vaccine;
 KW diphtheria; tetanus; polio; Haemophilus influenza b infection; therapy;
 KW immune response; maltose binding protein; MBP; fusion protein.
 XX OS Chimeric - Bordetella pertussis.
 OS Chimeric - Unidentified.
 XX PN WO200200695-A2.
 XX PD 03-JAN-2002.

XX PF 29-JUN-2001; 2001WO-NL00493.
 XX PR 30-JUN-2000; 2000EP-0202309.
 XX PA (NEME-) NEDERLANDEN MIN WELZIJN.
 XX PI Mooi FR;
 XX DR WPI; 2002-139897/18.
 XX PT New polypeptides derived from Bordetella pertussis pertactin, useful as
 PT a vaccine against infections caused by Bordetella strains, and other
 PT infectious diseases of mammals, e.g. diphtheria, tetanus, or polio -
 XX PS Example 1; Fig 4b; 52pp; English.
 XX CC The invention relates to polypeptides derived from Bordetella pertussis
 CC pertactin (Prt1). The polypeptide is useful in the preparation of
 CC vaccines against B. pertussis, B. parapertussis, B. bronchiseptica and
 CC other infectious diseases of mammals including diphtheria, tetanus, polio
 CC and infections caused by Haemophilus influenza b. The polypeptide is
 CC especially useful for eliciting an immune response against Bordetella sp.
 CC Antibodies against the polypeptide may be used for pharmaceutical and/or
 CC diagnostic purposes, particularly for treating or preventing infections
 CC caused by Bordetella pertussis or Bordetella parapertussis. The present
 CC sequence is maltose binding protein (MBP)-pertactin region 1 fusion
 CC protein used in the exemplification of the invention.
 XX SQ Sequence 36 AA;
 Query Match 100.0%; Score 31; DB 23; Length 36;
 Best Local Similarity 100.0%; Pred. No. 53;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGFPG 5
 DB 20 GGFPG 24
 DE 18-APR-2002 (first entry)
 AC AAE17174;
 AC AAE17174;
 DE Maltose binding protein-pertactin region 1 fusion protein, MBP-Prt5.
 DE Pertactin; prt1; antibacterial; immunostimulant; antimicrobial; vaccine;
 KW diphtheria; tetanus; polio; Haemophilus influenza b infection; therapy;
 KW immune response; maltose binding protein; MBP; fusion protein.
 XX OS Chimeric - Bordetella pertussis.
 OS Chimeric - Unidentified.
 XX PN WO200200695-A2.
 XX PD 03-JAN-2002.
 XX PF 29-JUN-2001; 2001WO-NL00493.
 XX PR 30-JUN-2000; 2000EP-0202309.
 XX PA (NEME-) NEDERLANDEN MIN WELZIJN.
 XX PI Mooi FR;
 XX DR WPI; 2002-139897/18.
 XX PT New polypeptides derived from Bordetella pertussis pertactin, useful as
 PT a vaccine against infections caused by Bordetella strains, and other

PT Infectious diseases of mammals, e.g. diphtheria, tetanus, or polio -
 XX Example 1; Fig 4b; 52pp; English.
 CC The invention relates to polypeptides derived from Bordetella pertussis
 CC pertactin (Prn1). The polypeptide is useful in the preparation of
 CC vaccines against B. pertussis, B. parapertussis, B. bronchiseptica and
 CC other infectious diseases of mammals including diphtheria, tetanus, polio
 CC and infections caused by Haemophilus influenza b. The polypeptide is
 CC especially useful for eliciting an immune response against Bordetella sp.
 CC diagnostic purposes, particularly for treating or preventing infections
 CC caused by Bordetella pertussis or Bordetella parapertussis. The present
 CC sequence is maltose binding protein (MBP)-pertactin region 1 fusion
 CC protein used in the exemplification of the invention.
 SQ Sequence 36 AA;
 Query Match 100.0%; Score 31; DB 23; Length 36;
 Best Local Similarity 100.0%; Pred. No. 53;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGFEP 5
 |||||
 DB 20 GGFEP 24
 RESULT 14
 ID AAE17171 standard; Protein; 41 AA.
 XX AAE17171;
 AC AAE17171;
 DT 18-APR-2002 (first entry)
 DE Maltose binding protein-pertactin region 1 fusion protein, MBP-Prn2.
 XX Pertactin; prn1; antibacterial; immunostimulant; antimicrobial; vaccine;
 KW diphtheria; tetanus; polio; Haemophilus influenza b infection; therapy;
 KM immune response; maltose binding protein; MBP; fusion protein.
 XX Chimeric - Bordetella pertussis.
 OS Chimeric - Unidentified.
 PN WO200200695-A2.
 PD 03-JAN-2002.
 PF 29-JUN-2001; 2001MO-NI00493.
 XX 30-JUN-2000; 2000EP-0202309.
 PR (NWE-) NEDERLANDEN MIN WELZIJN.
 PA (NWE-) NEDERLANDEN MIN WELZIJN.
 PI Mooi FR;
 XX WPI; 2002-139897/18.
 DR New polypeptides derived from Bordetella pertussis pertactin, useful as
 PT a vaccine against infections caused by Bordetella strains, and other
 PT infectious diseases of mammals, e.g. diphtheria, tetanus, or polio -
 XX Example 1; Fig 4b; 52pp; English.
 XX The invention relates to polypeptides derived from Bordetella pertussis
 CC pertactin (Prn1). The polypeptide is useful in the preparation of
 CC vaccines against B. pertussis, B. parapertussis, B. bronchiseptica and
 CC other infectious diseases of mammals including diphtheria, tetanus, polio
 CC and infections caused by Haemophilus influenza b. The polypeptide is
 CC especially useful for eliciting an immune response against Bordetella sp.
 CC antibodies against the polypeptide may be used for pharmaceutical and/or
 CC diagnostic purposes, particularly for treating or preventing infections
 CC caused by Bordetella pertussis or Bordetella parapertussis. The present

CC sequence is maltose binding protein (MBP)-pertactin region 1 fusion
 CC protein used in the exemplification of the invention.
 SQ Sequence 41 AA;
 Query Match 100.0%; Score 31; DB 23; Length 41;
 Best Local Similarity 100.0%; Pred. No. 60;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGFEP 5
 |||||
 DB 20 GGFEP 24
 RESULT 15
 ID AAE17153 standard; Protein; 45 AA.
 XX AAE17153;
 AC AAE17153;
 DT 18-APR-2002 (first entry)
 DE F647 monoclonal antibody specific for region 1 of pertactin.
 XX Pertactin; prn1; antibacterial; immunostimulant; antimicrobial; vaccine;
 KW diphtheria; tetanus; polio; Haemophilus influenza b infection; therapy;
 KM immune response; antibody.
 XX Unidentified.
 OS Unidentified.
 FH Key Location/Qualifiers
 FT Region 31..40
 /note="Epitope"
 PN WO200200695-A2.
 PD 03-JAN-2002.
 PF 29-JUN-2001; 2001MO-NI00493.
 XX 30-JUN-2000; 2000EP-0202309.
 PR (NWE-) NEDERLANDEN MIN WELZIJN.
 PA (NWE-) NEDERLANDEN MIN WELZIJN.
 PI Mooi FR;
 XX WPI; 2002-139897/18.
 DR New polypeptides derived from Bordetella pertussis pertactin, useful as
 PT a vaccine against infections caused by Bordetella strains, and other
 PT infectious diseases of mammals, e.g. diphtheria, tetanus, or polio -
 XX Example 1; Fig 2a; 52pp; English.
 XX The invention relates to polypeptides derived from Bordetella pertussis
 CC pertactin (Prn1). The polypeptide is useful in the preparation of
 CC vaccines against B. pertussis, B. parapertussis, B. bronchiseptica and
 CC other infectious diseases of mammals including diphtheria, tetanus, polio
 CC and infections caused by Haemophilus influenza b. The polypeptide is
 CC especially useful for eliciting an immune response against Bordetella sp.
 CC antibodies against the polypeptide may be used for pharmaceutical and/or
 CC diagnostic purposes, particularly for treating or preventing infections
 CC caused by Bordetella pertussis or Bordetella parapertussis. The present
 CC sequence is a monoclonal antibody specific to region 1 of pertactin
 CC used for epitope mapping and vaccination.
 SQ Sequence 45 AA;
 Query Match 100.0%; Score 31; DB 23; Length 45;
 Best Local Similarity 100.0%; Pred. No. 65;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGFEP 5

Db 26 GFGP 30

Search completed: March 17, 2003, 12:08:56
Job time : 32.3333 secs

GenCore version 5.1.4.P5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 17, 2003, 12:06:59 ; Search time 12 Seconds
(without alignments)
12.260 Million cell updates/sec

Title: SEQ25-FG
Perfect score: 31
Sequence: 1 g9fgp 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PTUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31	100.0	129	1 US-08-462-949-29	Sequence 29, Appl
2	31	100.0	129	1 US-08-023-764B-29	Sequence 29, Appl
3	31	100.0	204	4 US-09-323-872A-13	Sequence 13, Appl
4	31	100.0	254	4 US-09-362-831-3	Sequence 3, Appl
5	31	100.0	733	4 US-09-073-587-3	Sequence 3, Appl
6	31	100.0	739	4 US-09-503-922-1	Sequence 1, Appl
7	31	100.0	796	4 US-08-817-800-2	Sequence 2, Appl
8	31	100.0	796	4 US-09-236-645-2	Sequence 2, Appl
9	31	100.0	907	2 US-09-010-928B-4	Sequence 4, Appl
10	31	100.0	910	4 US-08-460-269C-2	Sequence 2, Appl
11	31	100.0	911	4 US-08-460-269C-4	Sequence 4, Appl
12	31	100.0	922	4 US-08-460-269C-6	Sequence 6, Appl
13	31	100.0	1186	2 US-08-861-664-8	Sequence 8, Appl
14	31	100.0	1186	2 US-08-386-001-8	Sequence 8, Appl
15	31	100.0	1186	4 US-09-323-833A-8	Sequence 8, Appl
16	31	100.0	1207	1 US-07-951-715A-7	Sequence 7, Appl
17	31	100.0	1207	3 US-08-459-448A-7	Sequence 7, Appl
18	31	100.0	1207	3 US-08-459-595A-7	Sequence 7, Appl
19	31	100.0	1207	3 US-08-459-504B-7	Sequence 7, Appl
20	31	100.0	1207	3 US-08-459-444-7	Sequence 7, Appl
21	31	100.0	1207	3 US-09-053-549-8	Sequence 8, Appl
22	31	100.0	1207	3 US-09-547-422-7	Sequence 8, Appl
23	31	100.0	1227	3 US-09-053-549-2	Sequence 2, Appl
24	31	100.0	3729	2 US-08-804-227C-4	Sequence 4, Appl
25	31	100.0	3816	4 US-09-428-517-3	Sequence 3, Appl
26	28	90.3	12	4 US-09-039-780A-108	Sequence 108, Appl
27	28	90.3	14	1 US-08-209-747-35	Sequence 35, Appl

28	28	90.3	14	1 US-08-458-298-35	Sequence 35, Appl
29	28	90.3	16	4 US-09-039-780A-107	Sequence 107, Appl
30	28	90.3	18	4 US-08-556-978B-52	Sequence 52, Appl
31	28	90.3	19	4 US-08-556-978B-9	Sequence 9, Appl
32	28	90.3	20	2 US-09-010-928B-7	Sequence 7, Appl
33	28	90.3	22	2 US-08-317-844B-8	Sequence 8, Appl
34	28	90.3	22	4 US-08-556-978B-46	Sequence 46, Appl
35	28	90.3	26	4 US-09-039-780A-13	Sequence 13, Appl
36	28	90.3	28	4 US-08-556-978B-58	Sequence 58, Appl
37	28	90.3	29	1 US-08-425-069-31	Sequence 31, Appl
38	28	90.3	29	1 US-08-209-747-42	Sequence 42, Appl
39	28	90.3	29	1 US-08-458-298-42	Sequence 42, Appl
40	28	90.3	29	2 US-08-317-844B-31	Sequence 31, Appl
41	28	90.3	32	1 US-08-425-069-34	Sequence 34, Appl
42	28	90.3	32	1 US-08-209-747-45	Sequence 45, Appl
43	28	90.3	32	1 US-08-458-298-45	Sequence 45, Appl
44	28	90.3	32	2 US-08-317-844B-34	Sequence 34, Appl
45	28	90.3	35	1 US-08-209-747-52	Sequence 52, Appl

ALIGNMENTS

RESULT 1
US-08-462-949-29
Sequence 29, Application US/08462949
Patent No. 5606022
GENERAL INFORMATION:
APPLICANT: Rasmussen, Beth Ann
TITLE OF INVENTION: Cloning and Identification of a Two Component Signal Transducing Regulatory System from Bacteroides Fragilis
TITLE OF INVENTION: Bacteroides Fragilis
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Darby & Darby P.C.
STREET: 805 Third Avenue
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462,949
CLASSIFICATION: 435
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/023,764
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Robinson, Joseph R.
REGISTRATION NUMBER: 33,448
REFERENCE/DOCKET NUMBER: 0646/18024-US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-527-7700
TELEFAX: 201-753-6237
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 129 amino acids
TYPE: amino acid
STRANDEDNESS: single
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-462-949-29
Query Match 100.0%; Score 31; DB 1; Length 129;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 5; Conservative 0; Mismatches 0; Gaps 0;

OY 1 GGF 5
 DB 49 GGF 53

RESULT 2

US-08-023-764B-29
 ; Sequence 29, Application US/08023764B
 ; Patent No. 5679540
 ; GENERAL INFORMATION:
 ; APPLICANT: Rasmussen, Beth Ann
 ; TITLE OF INVENTION: Cloning and Identification of a Two
 ; TITLE OF INVENTION: Component Signal Transducing Regulatory System from
 ; TITLE OF INVENTION: Bacteroides fragilis
 ; NUMBER OF SEQUENCES: 39
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: American Cyanamid Company
 ; STREET: One Cyanamid Plaza
 ; CITY: Wayne
 ; STATE: New Jersey
 ; COUNTRY: United States
 ; ZIP: 07470-8426
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: IBM PC compatible
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/023,764B
 ; FILING DATE: 26-FEB-1993
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Barnhard, Elizabeth M.
 ; REGISTRATION NUMBER: 31,088
 ; REFERENCE/DOCKET NUMBER: 31,658-00
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (201)831-3246
 ; TELEFAX: (201)831-3305
 ; INFORMATION FOR SEQ ID NO: 29:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 129 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; HYPOTHEICAL: NO
 ; ANTI-SENSE: NO
 ; US-08-023-764B-29

Query Match 100.0%; Score 31; DB 1; Length 129;
 Best Local Similarity 100.0%; Pred. No. 63;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 GGF 5
 DB 49 GGF 53

RESULT 3

US-09-323-872A-13
 ; Sequence 13, Application US/09323872A
 ; Patent No. 6395539
 ; GENERAL INFORMATION:
 ; APPLICANT: Coschigno, Peter
 ; TITLE OF INVENTION: Compositions and Methods for Bioremediation
 ; FILE REFERENCE: OHU-03640
 ; CURRENT APPLICATION NUMBER: US/09/323,872A
 ; CURRENT FILING DATE: 2001-06-15
 ; PRIOR APPLICATION NUMBER: 09/072,433
 ; PRIOR FILING DATE: 1998-05-04
 ; NUMBER OF SEQ ID NOS: 58
 ; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 13
 ; LENGTH: 204
 ; TYPE: PRT
 ; ORGANISM: Rhizobium meliloti
 ; US-09-323-872A-13

Query Match 100.0%; Score 31; DB 4; Length 204;
 Best Local Similarity 100.0%; Pred. No. 94;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGF 5
 DB 199 GGF 203

RESULT 4

US-09-362-831-3
 ; Sequence 3, Application US/09362831
 ; Patent No. 6306400
 ; GENERAL INFORMATION:
 ; APPLICANT: BURLOT et al.
 ; TITLE OF INVENTION: AVIAN RECOMBINANT LIVE VACCINE USING, AS VECTOR, THE
 ; FILE REFERENCE: 454313-2520
 ; CURRENT APPLICATION NUMBER: US/09/362,831
 ; CURRENT FILING DATE: 1999-07-28
 ; NUMBER OF SEQ ID NOS: 28
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 3
 ; LENGTH: 254
 ; TYPE: PRT
 ; ORGANISM: Infectious Laryngotracheitis Virus
 ; US-09-362-831-3

Query Match 100.0%; Score 31; DB 4; Length 254;
 Best Local Similarity 100.0%; Pred. No. 1,1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 GGF 5
 DB 184 GGF 188

RESULT 5

US-09-073-587-3
 ; Sequence 3, Application US/09073587
 ; Patent No. 6268552
 ; GENERAL INFORMATION:
 ; APPLICANT: LI, YI
 ; TITLE OF INVENTION: Transgenic Seedless Plants
 ; NUMBER OF SEQUENCES: 7
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Greenlee, Wanner and Sullivan, P.C.
 ; STREET: 5370 Manhattan Circle, Suite 201
 ; CITY: Boulder
 ; STATE: Colorado
 ; COUNTRY: US
 ; ZIP: 80303
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/073,587
 ; FILING DATE: 06-MAY-1998
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 60/045,725
 ; FILING DATE: 06-MAY-1997
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Ferber, Donna M.
 ; REGISTRATION NUMBER: 33,878

REFERENCE/DOCKET NUMBER: 4-97
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 499-8080
TELEFAX: (303) 499-8089
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 733 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-073-587-3

Query Match 100.0%; Score 31; DB 4; Length 733;
Best local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFGP 5
Db 417 GGFGP 421

RESULT 6
US-09-503-922-1
Sequence 1, Application US/09503922
Patent No. 6410706
GENERAL INFORMATION:
APPLICANT: PAI, Hyun-Sook
APPLICANT: LIO, Jang-Ryol
APPLICANT: CHO, Hye-Sun
APPLICANT: KIM, Youn-Sung
TITLE OF INVENTION: A NOVEL CHITIN-BINDING RECEPTOR KINASE AND THE GENE
FILE REFERENCE: seq
CURRENT APPLICATION NUMBER: US/09/503,922
CURRENT FILING DATE: 2000-02-14
NUMBER OF SEQ ID NOS: 2
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 1
LENGTH: 739
TYPE: PRT
ORGANISM: Nicotiana tabacum
US-09-503-922-1

Query Match 100.0%; Score 31; DB 4; Length 739;
Best local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFGP 5
Db 430 GGFGP 434

RESULT 7
US-08-817-900-2
Sequence 2, Application US/08817900
Patent No. 5919648
GENERAL INFORMATION:
APPLICANT: BURR, KEITH WILLIAM
APPLICANT: RAMSDEN, MARTIN
APPLICANT: ILLING, GRAHAM TIMOTHY
APPLICANT: HARRISON, LESLIE ANN
APPLICANT: MAISHMAN, NICHOLAS JOHN
APPLICANT: SPENCE, DAVID WILSON
TITLE OF INVENTION: INDUSTRIAL ENZYMES
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHIVE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/817,900
FILING DATE: 25-APR-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9423212.1
FILING DATE: 17-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: CRAWFORD, ARTHUR R.
REGISTRATION NUMBER: 25,327
REFERENCE/DOCKET NUMBER: 1430-157
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 796 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-817-900-2

Query Match 100.0%; Score 31; DB 2; Length 796;
Best local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFGP 5
Db 300 GGFGP 304

RESULT 8
US-09-236-645-2
Sequence 2, Application US/09236645
Patent No. 6297032
GENERAL INFORMATION:
APPLICANT: BURR, KEITH WILLIAM
APPLICANT: RAMSDEN, MARTIN
APPLICANT: ILLING, GRAHAM TIMOTHY
APPLICANT: HARRISON, LESLIE ANN
APPLICANT: MAISHMAN, NICHOLAS JOHN
APPLICANT: SPENCE, DAVID WILSON
TITLE OF INVENTION: INDUSTRIAL ENZYMES
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHIVE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/236,645
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/817,900
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: CRAWFORD, ARTHUR R.
REGISTRATION NUMBER: 25,327
REFERENCE/DOCKET NUMBER: 1430-157

TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 796 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-236-645-2

Query Match 100.0%; Score 31; DB 4; Length 796;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFGP 5
|||||
DB 300 GGFGP 304

RESULT 9
US-09-010-928B-4

Sequence 4, Application US/09010928B
Patent No. 5994099
GENERAL INFORMATION:
APPLICANT: Lewis, Randolph V
TITLE OF INVENTION: EXTREMELY ELASTIC SPIDER SILK PROTEIN AND DNA
TITLE OF INVENTION: CODING THEREFOR
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESSES:
ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
STREET: 8110 GATEHOUSE RD, SUITE 500E
CITY: FALLS CHURCH
STATE: VIRGINIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 22042

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/010,928B
FILING DATE: 22-JAN-1998
CLASSIFICATION: A35
ATTORNEY/AGENT INFORMATION:
NAME: Murphy Jr., Gerald M
REGISTRATION NUMBER: 28977
REFERENCE/DOCKET NUMBER: 1447-109P
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 907 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-010-928B-4

Query Match 100.0%; Score 31; DB 2; Length 907;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFGP 5
|||||
DB 128 GGFGP 132

RESULT 10
US-08-460-269C-2

Sequence 2, Application US/08460269C
Patent No. 6197548
GENERAL INFORMATION:
APPLICANT: CLARE, JEFFREY J.

ROMANOS, MICHAEL A.
TITLE OF INVENTION: EXPRESSION OF HETEROLOGOUS PROTEIN IN YEAST
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Millen, White, Zelano & Branigan, P.C.
STREET: 2200 Clarendon Blvd., Suite 1400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22201

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,269C
FILING DATE: 02-Jun-1995
ATTORNEY/AGENT INFORMATION:
NAME: Lebovitz, Richard M.
REGISTRATION NUMBER: 37,067
REFERENCE/DOCKET NUMBER: Popov-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 243-6333
TELEFAX: (703) 243-6410
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 910 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-08-460-269C-2

Query Match 100.0%; Score 31; DB 4; Length 910;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFGP 5
|||||
DB 281 GGFGP 285

RESULT 11
US-08-460-269C-4

Sequence 4, Application US/08460269C
Patent No. 6197548
GENERAL INFORMATION:
APPLICANT: CLARE, JEFFREY J.
ROMANOS, MICHAEL A.
TITLE OF INVENTION: EXPRESSION OF HETEROLOGOUS PROTEIN IN YEAST
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Millen, White, Zelano & Branigan, P.C.
STREET: 2200 Clarendon Blvd., Suite 1400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22201

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,269C
FILING DATE: 02-Jun-1995
ATTORNEY/AGENT INFORMATION:
NAME: Lebovitz, Richard M.
REGISTRATION NUMBER: 37,067
REFERENCE/DOCKET NUMBER: Popov-2

US-08-460-269C-4
Sequence 4, Application US/08460269C
Patent No. 6197548
GENERAL INFORMATION:
APPLICANT: CLARE, JEFFREY J.
ROMANOS, MICHAEL A.
TITLE OF INVENTION: EXPRESSION OF HETEROLOGOUS PROTEIN IN YEAST
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Millen, White, Zelano & Branigan, P.C.
STREET: 2200 Clarendon Blvd., Suite 1400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22201
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,269C
FILING DATE: 02-Jun-1995
ATTORNEY/AGENT INFORMATION:
NAME: Lebovitz, Richard M.
REGISTRATION NUMBER: 37,067
REFERENCE/DOCKET NUMBER: Popov-2

TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 243-6333
TELEFAX: (703) 243-6410
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 911 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-08-460-269C-4

Query Match 100.0%; Score 31; DB 4; Length 911;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFSP 5
DB 276 GGFSP 280

RESULT 12
US-08-460-269C-6
Sequence 6, Application US/08460269C
Patent No. 6197548
GENERAL INFORMATION:
APPLICANT: CLARE, JEFFREY J.
ROMANOS, MICHAEL A.
TITLE OF INVENTION: EXPRESSION OF HETEROLOGOUS PROTEIN IN YEAST
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESSES:
ADDRESSEE: MILLER, White, Zelano & Branigan, P.C.
STREET: 2200 Clarendon Blvd., Suite 1400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22201
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,269C
FILING DATE: 02-Jun-1995
ATTORNEY/AGENT INFORMATION:
NAME: Lebovitz, Richard M.
REGISTRATION NUMBER: 37,067
REFERENCE/DOCKET NUMBER: Popov-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 243-6333
TELEFAX: (703) 243-6410
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 922 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-08-460-269C-6

Query Match 100.0%; Score 31; DB 4; Length 922;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFSP 5
DB 281 GGFSP 285

RESULT 13
US-08-861-464-8

Sequence 8, Application US/08861464
Patent No. 5874210
GENERAL INFORMATION:
APPLICANT: Guarente, Leonard P.
APPLICANT: Austriaco Jr., Nicanor
APPLICANT: Kennedy, Brian
TITLE OF INVENTION: Genes Determining Cellular Senescence
TITLE OF INVENTION: In Yeast
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Millitia Drive
CITY: Lexington
STATE: MA
COUNTRY: USA
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/861,464
FILING DATE: 22-MAY-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/396,001
FILING DATE: 28-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/09351
FILING DATE: 15-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/107,408
FILING DATE: 16-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: MIT-6408A22
TELECOMMUNICATION INFORMATION:
TELEPHONE: 781-861-6240
TELEFAX: 781-861-9540
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1186 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-861-464-8

Query Match 100.0%; Score 31; DB 2; Length 1186;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFSP 5
DB 238 GGFSP 242

RESULT 14
US-08-396-001-8
Sequence 8, Application US/08396001
Patent No. 5919618
GENERAL INFORMATION:
APPLICANT: Guarente, Leonard P.
APPLICANT: Austriaco Jr., Nicanor
APPLICANT: Claus, James
APPLICANT: Cole, Francesca
APPLICANT: Kennedy, Brian
TITLE OF INVENTION: Genes Determining Cellular Senescence In Yeast
TITLE OF INVENTION: Yeast
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.

DB 238 GGFGP 242
 Search completed: March 17, 2003, 12:11:51
 Job time: 13 secs

STREET: Two Millitia Drive
 CITY: Lexington
 STATE: MA
 COUNTRY: USA
 ZIP: 02173
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/396,001
 FILING DATE: 28-FEB-1995
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Granahan, Patricia
 REGISTRATION NUMBER: 32,227
 REFERENCE/DOCKET NUMBER: MIT-6408A2
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-861-6240
 TELEFAX: 617-861-9540
 INFORMATION FOR SEQ ID NO: 8:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1186 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-396-001-8

Query Match 100.0%; Score 31; DB 2; Length 1186;
 Best Local Similarity 100.0%; Pred. No. 4,4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFGP 5
 DB 238 GGFGP 242

RESULT 15
 US-09-323-433A-8
 ; Sequence 8, Application US/09323433A
 ; Patent No. 6218512
 ; GENERAL INFORMATION:
 ; APPLICANT: Guarente, Leonard P.
 ; APPLICANT: Austriaco Jr., Nicanor
 ; APPLICANT: Claus, James J.
 ; APPLICANT: Cole, Francesca
 ; APPLICANT: Kennedy, Brian
 ; TITLE OF INVENTION: GENES DETERMINING CELLULAR SENESENCE IN
 ; FILE REFERENCE: 0050,1491-003
 ; CURRENT APPLICATION NUMBER: US/09/323,433A
 ; CURRENT FILING DATE: 1999-06-01
 ; PRIOR APPLICATION NUMBER: US 08/396,001
 ; PRIOR FILING DATE: 1995-02-28
 ; PRIOR APPLICATION NUMBER: PCT/US94/09351
 ; PRIOR FILING DATE: 1994-08-15
 ; PRIOR APPLICATION NUMBER: US 08/107,408
 ; PRIOR FILING DATE: 1993-08-16
 ; NUMBER OF SEQ ID NOS: 48
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 8
 ; LENGTH: 1186
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-323-433A-8

Query Match 100.0%; Score 31; DB 4; Length 1186;
 Best Local Similarity 100.0%; Pred. No. 4,4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 GGFGP 5
 DB 238 GGFGP 242

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 17, 2003, 12:10:29 ; Search time 11 Seconds
(without alignments)
20.951 Million cell updates/sec

Title: SEQ25-FG
Perfect score: 31
Sequence: 1 gfgfp 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 199416 seqs, 46092074 residues
Total number of hits satisfying chosen parameters: 199416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
- 2: /cgn2_6/ptodata/1/pubpaa/PCY_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
- 6: /cgn2_6/ptodata/1/pubpaa/PCY07_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
- 12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
- 13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
- 14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result, being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	31	100.0	46 9 US-09-855-754-8	Sequence 8, Appl1
2	31	100.0	51 9 US-09-855-754-7	Sequence 7, Appl1
3	31	100.0	51 9 US-09-855-754-13	Sequence 13, Appl1
4	31	100.0	56 9 US-09-855-754-9	Sequence 9, Appl1
5	31	100.0	56 9 US-09-855-754-10	Sequence 10, Appl1
6	31	100.0	56 9 US-09-855-754-12	Sequence 12, Appl1
7	31	100.0	61 9 US-09-855-754-11	Sequence 11, Appl1
8	31	100.0	137 9 US-09-738-626-6472	Sequence 6472, Ap
9	31	100.0	157 9 US-09-252-088-10	Sequence 10, Appl1
10	31	100.0	183 9 US-10-101-464A-656	Sequence 656, Appl
11	31	100.0	229 9 US-09-738-626-3540	Sequence 3540, Ap
12	31	100.0	407 10 US-09-815-242-11180	Sequence 11180, A
13	31	100.0	466 9 US-09-738-626-4630	Sequence 4630, Ap
14	31	100.0	645 10 US-09-919-172-41	Sequence 41, Appl
15	31	100.0	658 9 US-09-801-220-4	Sequence 4, Appl1
16	31	100.0	798 9 US-10-101-464A-926	Sequence 926, App
17	31	100.0	803 9 US-09-801-220-2	Sequence 2, Appl1
18	31	100.0	910 9 US-09-855-754-5	Sequence 5, Appl1
19	31	100.0	911 9 US-09-855-754-4	Sequence 4, Appl1

20	31	100.0	922 9 US-09-855-754-6	Sequence 6, Appl1
21	31	100.0	939 10 US-09-854-845-16	Sequence 16, Appl
22	31	100.0	954 10 US-09-854-845-14	Sequence 14, Appl
23	31	100.0	1034 10 US-09-854-845-6	Sequence 6, Appl1
24	31	100.0	1049 10 US-09-854-845-2	Sequence 2, Appl1
25	31	100.0	1078 10 US-09-854-845-8	Sequence 8, Appl1
26	31	100.0	1093 10 US-09-854-845-4	Sequence 4, Appl1
27	31	100.0	1136 10 US-09-854-845-12	Sequence 12, Appl
28	31	100.0	1151 10 US-09-854-845-10	Sequence 10, Appl
29	31	100.0	1186 10 US-09-826-752-8	Sequence 8, Appl1
30	31	100.0	1207 9 US-09-988-462-7	Sequence 7, Appl1
31	31	100.0	3816 9 US-09-808-880-3	Sequence 3, Appl1
32	28	90.3	39 8 US-08-424-508-405	Sequence 405, Appl
33	28	90.3	47 10 US-09-040-518-3	Sequence 3, Appl1
34	28	90.3	60 10 US-09-864-761-47716	Sequence 4, Appl1
35	28	90.3	119 10 US-09-861-597-9	Sequence 47716, A
36	28	90.3	123 10 US-09-867-550-516	Sequence 9, Appl1
37	28	90.3	147 10 US-09-864-761-46558	Sequence 516, App
38	28	90.3	160 9 US-09-854-133-435	Sequence 46558, A
39	28	90.3	160 10 US-09-738-973-435	Sequence 435, App
40	28	90.3	174 10 US-09-818-066-60	Sequence 60, Appl
41	28	90.3	307 10 US-09-036-613-7	Sequence 7, Appl1
42	28	90.3	334 9 US-09-738-626-3720	Sequence 3720, Ap
43	28	90.3	527 9 US-09-712-363-156	Sequence 156, App
44	28	90.3	529 10 US-09-861-597-2	Sequence 2, Appl1
45	28	90.3	594 9 US-10-028-072-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-09-855-754-8
Sequence 8, Application US/09855754
Publication No. US20020192237A1
GENERAL INFORMATION:
APPLICANT: BOURSAX-EBUDE, CAROLINE
APPLICANT: GUISSO-MACLOUF, NICOLE
TITLE OF INVENTION: POLYPEPTIDES CONTRAINING POLYMORPHISMS OF THE REPEATED
TITLE OF INVENTION: REGIONS OF PERTACTIN IN BORDETTELLA PERTUSSIS,
TITLE OF INVENTION: BORDETTELLA PARAPERTUSSIS, AND BORDETTELLA
TITLE OF INVENTION: BRONCHISEPTICA, THEIR USE IN DIAGNOSTICS, AND IN
FILE REFERENCE: 03495-0206-00000
CURRENT APPLICATION NUMBER: US/09/855,754
CURRENT FILING DATE: 2001-09-10
PRIOR APPLICATION NUMBER: 60/206,969
PRIOR FILING DATE: 2000-05-25
NUMBER OF SEQ ID NOS: 24
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 8
LENGTH: 46
TYPE: PRT
ORGANISM: Bordetella bronchiseptica
US-09-855-754-8
Query Match
Best Local Similarity 100.0%: Score 31; DB 9; Length 46;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 GGFSP 5
Db 23 GGFSP 27
RESULT 2
US-09-855-754-7
Sequence 7, Application US/09855754
Publication No. US20020192237A1
GENERAL INFORMATION:
APPLICANT: BOURSAX-EBUDE, CAROLINE
APPLICANT: GUISSO-MACLOUF, NICOLE
TITLE OF INVENTION: POLYPEPTIDES CONTRAINING POLYMORPHISMS OF THE REPEATED

```
;; TITLE OF INVENTION: REGIONS OF PERTACTIN IN BORDETELLA PERTUSSIS,  
;; TITLE OF INVENTION: BORDETELLA PARAPERTUSSIS, AND BORDETELLA  
;; TITLE OF INVENTION: BRONCHISEPTICA, THEIR USE IN DIAGNOSTICS, AND IN  
;; FILE REFERENCE: 03495-0206-00000  
;; CURRENT APPLICATION NUMBER: US/09/855,754  
;; CURRENT FILING DATE: 2001-09-10  
;; PRIOR APPLICATION NUMBER: 60/206,969  
;; PRIOR FILING DATE: 2000-05-25  
;; NUMBER OF SEQ ID NOS: 24  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO 7  
;; LENGTH: 51  
;; TYPE: PRT  
;; ORGANISM: Bordetella bronchiseptica  
US-09-855-754-7
```

```
Query Match  
Best Local Similarity 100.0%; Score 31; DB 9; Length 51;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 GGFGP 5  
Db 28 GGFGP 32
```

```
RESULT 3  
US-09-855-754-13  
;; Sequence 13, Application US/09855754  
;; Publication No. US20020192237A1  
;; GENERAL INFORMATION:  
;; APPLICANT: BOURSAX-UEDE, CAROLINE  
;; APPLICANT: GUISSO-MACLOUF, NICOLE  
;; TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED  
;; TITLE OF INVENTION: REGIONS OF PERTACTIN IN BORDETELLA PERTUSSIS,  
;; TITLE OF INVENTION: BORDETELLA PARAPERTUSSIS, AND BORDETELLA  
;; TITLE OF INVENTION: BRONCHISEPTICA, THEIR USE IN DIAGNOSTICS, AND IN  
;; FILE REFERENCE: 03495-0206-00000  
;; CURRENT APPLICATION NUMBER: US/09/855,754  
;; CURRENT FILING DATE: 2001-09-10  
;; PRIOR APPLICATION NUMBER: 60/206,969  
;; PRIOR FILING DATE: 2000-05-25  
;; NUMBER OF SEQ ID NOS: 24  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO 13  
;; LENGTH: 51  
;; TYPE: PRT  
;; ORGANISM: Bordetella bronchiseptica  
US-09-855-754-13
```

```
Query Match  
Best Local Similarity 100.0%; Score 31; DB 9; Length 51;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 GGFGP 5  
Db 23 GGFGP 27
```

```
RESULT 4  
US-09-855-754-9  
;; Sequence 9, Application US/09855754  
;; Publication No. US20020192237A1  
;; GENERAL INFORMATION:  
;; APPLICANT: BOURSAX-UEDE, CAROLINE  
;; APPLICANT: GUISSO-MACLOUF, NICOLE  
;; TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED  
;; TITLE OF INVENTION: REGIONS OF PERTACTIN IN BORDETELLA PERTUSSIS,  
;; TITLE OF INVENTION: BORDETELLA PARAPERTUSSIS, AND BORDETELLA  
;; TITLE OF INVENTION: BRONCHISEPTICA, THEIR USE IN DIAGNOSTICS, AND IN  
;; FILE REFERENCE: 03495-0206-00000
```

```
;; CURRENT APPLICATION NUMBER: US/09/855,754  
;; CURRENT FILING DATE: 2001-09-10  
;; PRIOR APPLICATION NUMBER: 60/206,969  
;; PRIOR FILING DATE: 2000-05-25  
;; NUMBER OF SEQ ID NOS: 24  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO 9  
;; LENGTH: 56  
;; TYPE: PRT  
;; ORGANISM: Bordetella bronchiseptica  
US-09-855-754-9
```

```
Query Match  
Best Local Similarity 100.0%; Score 31; DB 9; Length 56;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 GGFGP 5  
Db 28 GGFGP 32
```

```
RESULT 5  
US-09-855-754-10  
;; Sequence 10, Application US/09855754  
;; Publication No. US20020192237A1  
;; GENERAL INFORMATION:  
;; APPLICANT: BOURSAX-UEDE, CAROLINE  
;; APPLICANT: GUISSO-MACLOUF, NICOLE  
;; TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED  
;; TITLE OF INVENTION: REGIONS OF PERTACTIN IN BORDETELLA PERTUSSIS,  
;; TITLE OF INVENTION: BORDETELLA PARAPERTUSSIS, AND BORDETELLA  
;; TITLE OF INVENTION: BRONCHISEPTICA, THEIR USE IN DIAGNOSTICS, AND IN  
;; FILE REFERENCE: 03495-0206-00000  
;; CURRENT APPLICATION NUMBER: US/09/855,754  
;; CURRENT FILING DATE: 2001-09-10  
;; PRIOR APPLICATION NUMBER: 60/206,969  
;; PRIOR FILING DATE: 2000-05-25  
;; NUMBER OF SEQ ID NOS: 24  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO 10  
;; LENGTH: 56  
;; TYPE: PRT  
;; ORGANISM: Bordetella bronchiseptica  
US-09-855-754-10
```

```
Query Match  
Best Local Similarity 100.0%; Score 31; DB 9; Length 56;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 GGFGP 5  
Db 28 GGFGP 32
```

```
RESULT 6  
US-09-855-754-12  
;; Sequence 12, Application US/09855754  
;; Publication No. US20020192237A1  
;; GENERAL INFORMATION:  
;; APPLICANT: BOURSAX-UEDE, CAROLINE  
;; APPLICANT: GUISSO-MACLOUF, NICOLE  
;; TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED  
;; TITLE OF INVENTION: REGIONS OF PERTACTIN IN BORDETELLA PERTUSSIS,  
;; TITLE OF INVENTION: BORDETELLA PARAPERTUSSIS, AND BORDETELLA  
;; TITLE OF INVENTION: BRONCHISEPTICA, THEIR USE IN DIAGNOSTICS, AND IN  
;; FILE REFERENCE: 03495-0206-00000  
;; CURRENT APPLICATION NUMBER: US/09/855,754  
;; CURRENT FILING DATE: 2001-09-10  
;; PRIOR APPLICATION NUMBER: 60/206,969  
;; PRIOR FILING DATE: 2000-05-25  
;; NUMBER OF SEQ ID NOS: 24
```


SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 12
LENGTH: 56
TYPE: PRT
ORGANISM: Bordetella bronchiseptica
US-09-855-754-12

Query Match
Best Local Similarity 100.0%; Score 31; DB 9; Length 56;
Pred. No. 19;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFGP 5
|||||
DB 23 GGFGP 27

RESULT 7
US-09-855-754-11
Sequence 11, Application US/09855754
Publication No. US20020192237A1
GENERAL INFORMATION:
APPLICANT: BOURSANX-EUDE, CAROLINE
APPLICANT: GUISSO-MACLOUF, NICOLE
TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED
TITLE OF INVENTION: REGIONS OF PERTACTIN IN BORDETELLA PERTUSSIS,
TITLE OF INVENTION: BORDETELLA PARAPERTUSSIS, AND BORDETELLA
TITLE OF INVENTION: BRONCHISEPTICA, THEIR USE IN DIAGNOSTICS, AND IN
FILE REFERENCE: 03495-0206-00000
CURRENT APPLICATION NUMBER: US/09/855,754
CURRENT FILING DATE: 2001-09-10
PRIOR APPLICATION NUMBER: 60/206,969
PRIOR FILING DATE: 2000-05-25
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 11
LENGTH: 61
TYPE: PRT
ORGANISM: Bordetella bronchiseptica
US-09-855-754-11

Query Match
Best Local Similarity 100.0%; Score 31; DB 9; Length 61;
Pred. No. 21;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFGP 5
|||||
DB 23 GGFGP 27

RESULT 8
US-09-738-626-6472
Sequence 6472, Application US/09738626
Publication No. US20020197605A1
GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIRO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAOKO
APPLICANT: SENOH, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07

PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 6472
LENGTH: 137
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-738-626-6472

Query Match
Best Local Similarity 100.0%; Score 31; DB 9; Length 137;
Pred. No. 43;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFGP 5
|||||
DB 115 GGFGP 119

RESULT 9
US-09-252-088-10
Sequence 10, Application US/09252088
Publication No. US20030031682A1
GENERAL INFORMATION:
APPLICANT: BRODEUR, Bernard R.
APPLICANT: RIOUX, Clément
APPLICANT: BOYER, Martine
APPLICANT: CHARLEBOIS, Isabelle
APPLICANT: HAMEL, Jose
APPLICANT: MARTIN, Denis
TITLE OF INVENTION: NOVEL GROUP B STREPTOCOCCUS ANTIGENS
FILE REFERENCE: 8331-9002
CURRENT APPLICATION NUMBER: US/09/252,088
CURRENT FILING DATE: 1999-02-18
EARLIER APPLICATION NUMBER: US/60/075,425
EARLIER FILING DATE: 1998-02-20
NUMBER OF SEQ ID NOS: 44
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 10
LENGTH: 157
TYPE: PRT
ORGANISM: group B streptococcus
US-09-252-088-10

Query Match
Best Local Similarity 100.0%; Score 31; DB 9; Length 157;
Pred. No. 49;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFGP 5
|||||
DB 67 GGFGP 71

RESULT 10
US-10-101-464A-656
Sequence 656, Application US/10101464A
Publication No. US20030046728A1
GENERAL INFORMATION:
APPLICANT: Strabala, Timothy
APPLICANT: Nieuwenhuizen, Nicolaas
APPLICANT: Higgins, Colleen M.
TITLE OF INVENTION: Compositions isolated from plant cells
TITLE OF INVENTION: and their use in the modification of plant cell signaling
FILE REFERENCE: 11000.1020C2
CURRENT APPLICATION NUMBER: US/10/101,464A
CURRENT FILING DATE: 2002-03-18
PRIOR APPLICATION NUMBER: 09/704,302
PRIOR FILING DATE: 2000-11-01
PRIOR APPLICATION NUMBER: 09/228,986
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/162,866
PRIOR FILING DATE: 1999-11-01
PRIOR APPLICATION NUMBER: PCT/US00/00724

```
;; PRIOR FILING DATE: 2000-01-11
;; NUMBER OF SEQ ID NOS: 989
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 656
;; LENGTH: 183
;; TYPE: PRT
;; ORGANISM: Pinus radiata
US-10-101-464A-656
```

```
Query Match
Best Local Similarity 100.0%; Score 31; DB 9; Length 183;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 GGFEP 5
DB 14 GGFEP 18
```

RESULT 11

```
US-09-738-626-3540
;; Sequence 3540, Application US/09738626
;; Publication No. US20020197605A1
;; GENERAL INFORMATION:
;; APPLICANT: NAKAGAWA, SATOSHI
;; APPLICANT: MIZOGUCHI, HIROSHI
;; APPLICANT: ANDO, SEIKO
;; APPLICANT: HAYASHI, MIKIRO
;; APPLICANT: OCHIAI, KEIKO
;; APPLICANT: YOKOI, HARUHIKO
;; APPLICANT: TATEISHI, NAKO
;; APPLICANT: SENOH, AKIHRO
;; APPLICANT: IKEDA, MASATO
;; APPLICANT: OZAKI, AKIO
;; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;; FILE REFERENCE: 249-125
;; CURRENT APPLICATION NUMBER: US/09/738,626
;; PRIOR FILING DATE: 2000-12-18
;; PRIOR APPLICATION NUMBER: JP 99/377484
;; PRIOR FILING DATE: 1999-12-16
;; PRIOR APPLICATION NUMBER: JP 00/159162
;; PRIOR FILING DATE: 2000-04-07
;; PRIOR APPLICATION NUMBER: JP 00/280988
;; PRIOR FILING DATE: 2000-08-03
;; NUMBER OF SEQ ID NOS: 7059
;; SOFTWARE: PatentIn ver. 3.0
;; SEQ ID NO 3540
;; LENGTH: 229
;; TYPE: PRT
;; ORGANISM: Corynebacterium glutamicum
US-09-738-626-3540
```

```
Query Match
Best Local Similarity 100.0%; Score 31; DB 9; Length 229;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 GGFEP 5
DB 67 GGFEP 71
```

RESULT 12

```
US-09-815-242-11180
;; Sequence 11180, Application US/09815242
;; Patent No. US20020061569A1
;; GENERAL INFORMATION:
;; APPLICANT: Haselbeck, Robert
;; APPLICANT: Ohlser, Kari L.
;; APPLICANT: Zyskind, Judith W.
;; APPLICANT: Wall, Daniel
;; APPLICANT: Tirawick, John D.
;; APPLICANT: Grant J.
;; APPLICANT: Yamamoto, Robert T.
;; APPLICANT: Xu, H. Howard
```

```
;; TITLE OF INVENTION: Identification of Essential Genes in
;; FILE REFERENCE: ELITRA.011A
;; CURRENT APPLICATION NUMBER: US/09/815,242
;; PRIOR FILING DATE: 2001-03-21
;; PRIOR APPLICATION NUMBER: 60/191,078
;; PRIOR FILING DATE: 2000-03-21
;; PRIOR APPLICATION NUMBER: 60/206,848
;; PRIOR FILING DATE: 2000-05-23
;; PRIOR APPLICATION NUMBER: 60/207,727
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: 60/242,578
;; PRIOR FILING DATE: 2000-10-23
;; PRIOR APPLICATION NUMBER: 60/253,625
;; PRIOR FILING DATE: 2000-11-27
;; PRIOR APPLICATION NUMBER: 60/257,931
;; PRIOR FILING DATE: 2000-12-22
;; PRIOR APPLICATION NUMBER: 60/269,308
;; PRIOR FILING DATE: 2001-02-16
;; NUMBER OF SEQ ID NOS: 14110
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 11180
;; LENGTH: 407
;; TYPE: PRT
;; ORGANISM: Haemophilus influenzae
US-09-815-242-11180
```

```
Query Match
Best Local Similarity 100.0%; Score 31; DB 10; Length 407;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 GGFEP 5
DB 360 GGFEP 364
```

RESULT 13

```
US-09-738-626-4630
;; Sequence 4630, Application US/09738626
;; Publication No. US20020197605A1
;; GENERAL INFORMATION:
;; APPLICANT: NAKAGAWA, SATOSHI
;; APPLICANT: MIZOGUCHI, HIROSHI
;; APPLICANT: ANDO, SEIKO
;; APPLICANT: HAYASHI, MIKIRO
;; APPLICANT: OCHIAI, KEIKO
;; APPLICANT: YOKOI, HARUHIKO
;; APPLICANT: TATEISHI, NAKO
;; APPLICANT: SENOH, AKIHRO
;; APPLICANT: IKEDA, MASATO
;; APPLICANT: OZAKI, AKIO
;; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;; FILE REFERENCE: 249-125
;; CURRENT APPLICATION NUMBER: US/09/738,626
;; PRIOR FILING DATE: 2000-12-18
;; PRIOR APPLICATION NUMBER: JP 99/377484
;; PRIOR FILING DATE: 1999-12-16
;; PRIOR APPLICATION NUMBER: JP 00/159162
;; PRIOR FILING DATE: 2000-04-07
;; PRIOR APPLICATION NUMBER: JP 00/280988
;; PRIOR FILING DATE: 2000-08-03
;; NUMBER OF SEQ ID NOS: 7059
;; SOFTWARE: PatentIn ver. 3.0
;; SEQ ID NO 4630
;; LENGTH: 496
;; TYPE: PRT
;; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4630
```

```
Query Match
Best Local Similarity 100.0%; Score 31; DB 9; Length 496;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

OY 1 GGFGP 5
 |||||
 Db 383 GGFGP 387

RESULT 14
 US-09-919-172-41
 ; Sequence 41, Application US/09919172
 ; Patent No. US20020119463A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Paris, Mary
 ; APPLICANT: Turner, Christopher M.
 ; TITLE OF INVENTION: PROSTATE CANCER MARKERS
 ; FILE REFERENCE: PA-0036 US
 ; CURRENT APPLICATION NUMBER: US/09/919,172
 ; CURRENT FILING DATE: 2001-07-30
 ; PRIOR APPLICATION NUMBER: 60/222,469
 ; PRIOR FILING DATE: 2000-07-28
 ; NUMBER OF SEQ ID NOS: 102
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 41
 ; LENGTH: 645
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; OTHER INFORMATION: Incyte ID No. US20020119463A1 1798379CD1
 US-09-919-172-41

Query Match 100.0%; Score 31; DB 10; Length 645;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

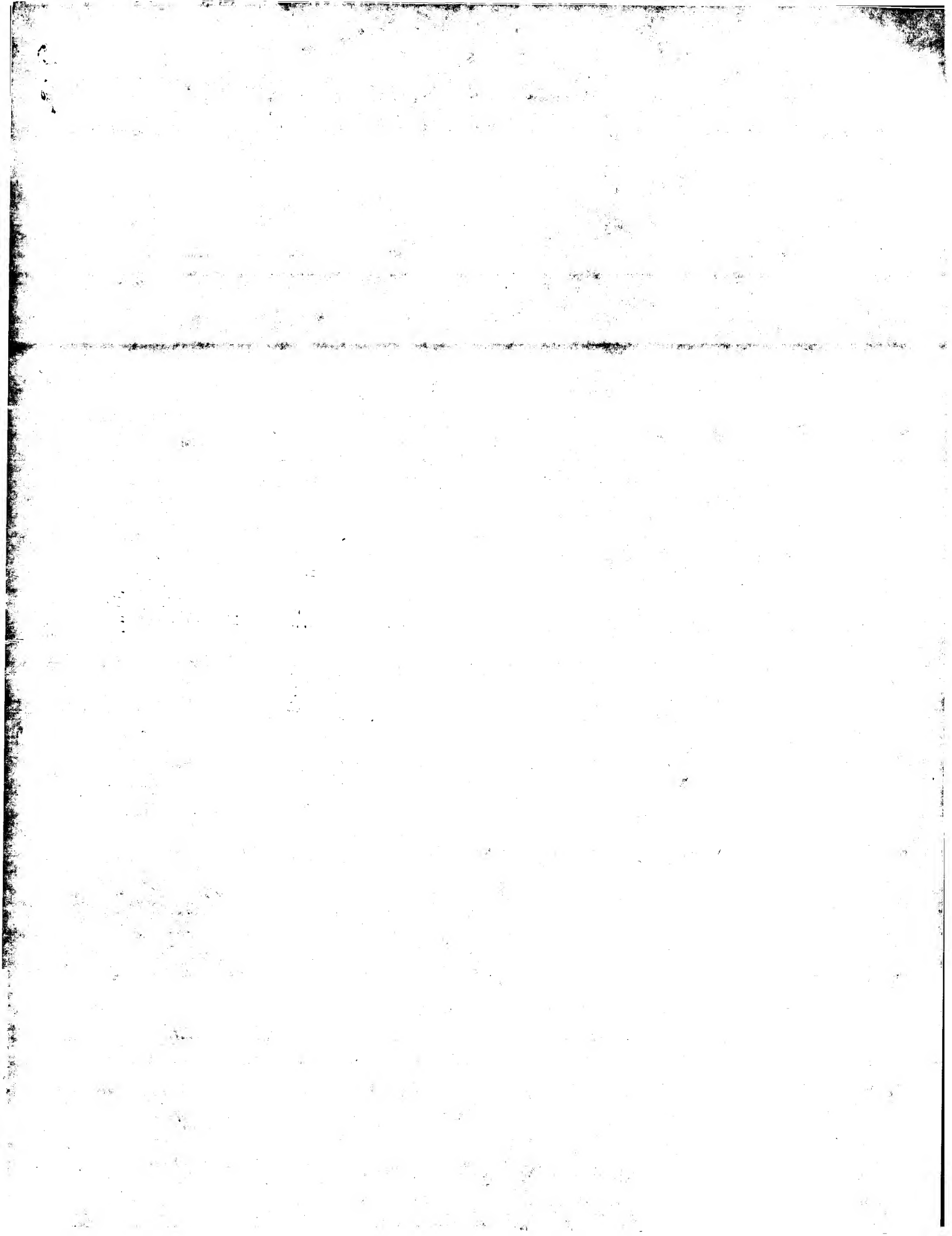
OY 1 GGFGP 5
 |||||
 Db 145 GGFGP 149

RESULT 15
 US-09-801-220-4
 ; Sequence 4, Application US/09801220
 ; Patent No. US20020173020A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Meyers, Rachel A.
 ; APPLICANT: Williamson, Mark
 ; TITLE OF INVENTION: 26886, A NOVEL CARBONIC ANHYDRASE
 ; TITLE OF INVENTION: FAMILY MEMBER AND USES THEREFOR
 ; FILE REFERENCE: 10448-026001
 ; CURRENT APPLICATION NUMBER: US/09/801,220
 ; CURRENT FILING DATE: 2001-03-07
 ; PRIOR APPLICATION NUMBER: US 60/187,456
 ; PRIOR FILING DATE: 2000-03-07
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 4
 ; LENGTH: 698
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: consensus sequence
 US-09-801-220-4

Query Match 100.0%; Score 31; DB 9; Length 698;
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFGP 5
 |||||
 Db 645 GGFGP 649

Search completed: March 17, 2003, 12:21:34
 Job time : 12 secs



GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 17, 2003, 12:06:28 ; Search time: 12.3333 Seconds
(without alignments)
38.973 Million cell updates/sec

Title: SEQ25-FG
Perfect score: 31
Sequence: 1 g9fgp 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	31	100.0	56	2 C72613	hypothetical prote
2	31	100.0	74	2 T33087	hypothetical prote
3	31	100.0	76	2 T09262	glycine-rich cell
4	31	100.0	130	2 D72780	hypothetical prote
5	31	100.0	169	2 B97483	hypothetical prote
6	31	100.0	204	2 B31227	nitrogen fixation
7	31	100.0	204	2 E95345	FixJ Transcription
8	31	100.0	238	2 C86656	glycerol uptake fa
9	31	100.0	269	2 D70660	probable aroE - My
10	31	100.0	280	1 C70784	probable dihem c
11	31	100.0	289	2 C87019	hypothetical prote
12	31	100.0	295	2 A87703	citrate lyase beta
13	31	100.0	299	2 G71088	hypothetical prote
14	31	100.0	326	2 F83802	hypothetical prote
15	31	100.0	348	2 D98326	probable zinc-bind
16	31	100.0	348	2 AB2957	zinc-binding dehyd
17	31	100.0	350	2 T29409	hypothetical prote
18	31	100.0	352	2 C95895	probable alcohol d
19	31	100.0	357	2 AH2756	hypothetical prote
20	31	100.0	359	2 D95876	conserved hypotet
21	31	100.0	367	2 AF3579	aminomethyltransfe
22	31	100.0	373	2 A83341	glycine cleavage s
23	31	100.0	374	1 LYYXLY	beta-lytic metallo
24	31	100.0	379	2 G97537	glycine cleavage s
25	31	100.0	381	2 H96731	hypothetical prote
26	31	100.0	389	2 T29410	hypothetical prote
27	31	100.0	394	2 T32964	hypothetical prote
28	31	100.0	394	2 C85064	hypothetical prote
29	31	100.0	395	2 T29411	hypothetical prote

30	31	100.0	396	2 T45033	hypothetical prote
31	31	100.0	407	2 C64167	hypothetical prote
32	31	100.0	424	2 F84658	probable protein k
33	31	100.0	427	2 T01903	hypothetical prote
34	31	100.0	430	2 S28707	hypothetical prote
35	31	100.0	462	2 C84473	probable protein k
36	31	100.0	469	2 A83356	hypothetical prote
37	31	100.0	475	2 H84137	hypothetical prote
38	31	100.0	518	2 S42091	rid(56) protein -
39	31	100.0	522	2 T28687	hypothetical prote
40	31	100.0	524	2 A69081	acetylactate synth
41	31	100.0	526	1 S33799	RNA-binding protei
42	31	100.0	528	2 G02127	tus-1-like protein
43	31	100.0	547	2 F82162	paraguat-inducible
44	31	100.0	561	2 T05545	pollen-specific pr
45	31	100.0	638	2 I53169	cytokeratin 2 - hu

ALIGNMENTS

RESULT 1

C72613
hypothetical protein APE051 - Aeropyrum pernix (strain K1)

C:Species: Aeropyrum pernix

C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999

C:Accession: C72613

R:Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Jin-no, K.; Ta
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.

DNA Res. 6, 83-101, 1999

A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aero

A:Reference number: A72450; MUID:99310339; PMID:10582966

A:Accession: C72613

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-56 <KAW>

A:Cross-references: DDBJ:AP000061; NID:q5104821; PIDN:BAAB0361.1; PID:dl044147; PID:g

A:Experimental source: strain K1

C:Genetics:

A:Gene: APE051

Query Match 100.0%; Score 31; DB 2; Length 56;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFEP 5
Db 42 GGFEP 46

RESULT 2

T33087
hypothetical protein R12E2.6 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

C:Accession: T33087

R:Goela, D.; Scheet, P.
submitted to the EMBL Data Library, May 1998

A:Description: The sequence of C. elegans cosmid R12E2.

A:Reference number: 221281

A:Accession: T33087

A:Status: preliminary; translated from CB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-74 <GOE>

A:Cross-references: EMBL:AR067219; PIDN:AACT17027.1; GSPDB:GN00019; CESP:R12E2.6

A:Experimental source: strain Bristol N2; clone R12E2

C:Genetics:

A:Gene: CESP:R12E2.6

A:Map position: 1

A:Introns: 43/2

Query Match 100.0%; Score 31; DB 2; Length 74;
Best Local Similarity 100.0%; Pred. No. 28;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGF 5
 Db 45 GGF 49

RESULT 3

T09262
 glycine-rich cell wall protein EMB31 - white spruce
 C:Species: Picea glauca (white spruce)
 C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 11-Jun-1999
 C:Accession: T09262
 R:Don, J.Z.; Dunstan, D.I.
 Submitted to the EMBL Data Library, June 1996
 A:Description: Gene expression during somatic embryogenesis.
 A:Reference number: Z16588
 A:Accession: T09262
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-76 <DON>
 A:Cross-references: EMBL:L47748; NID:g1350525; PID:g1350526
 C:Genetics:
 A:Gene: EMB31

Query Match 100.0%; Score 31; DB 2; Length 76;
 Best Local Similarity 100.0%; Pred. No. 29;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGF 5
 Db 26 GGF 30

RESULT 4

D72780
 hypothetical protein APE0229 - Aeropyrum pernix (strain K1)
 C:Species: Aeropyrum pernix
 C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
 C:Accession: D72780
 R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takai
 dna, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
 dna Res. 6, 83-101, 1999
 A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
 A:Reference number: A72450; MUID:99310339; PMID:10382966
 A:Accession: D72780
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-130 <KAN>
 A:Cross-references: DDBJ:AP000058; NID:g5103388; PIDN:BAAT9142.1; PID:g5103621
 A:Experimental source: strain K1
 C:Genetics:
 A:Gene: APE0229
 C:Superfamily: Aeropyrum pernix hypothetical protein APE0229

Query Match 100.0%; Score 31; DB 2; Length 130;
 Best Local Similarity 100.0%; Pred. No. 49;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGF 5
 Db 61 GGF 65

RESULT 5

B97483
 hypothetical protein AGR_C_1859 [imported] - Agrobacterium tumefaciens (strain C58, Cere
 C:Species: Agrobacterium tumefaciens
 C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002
 C:Accession: B97483
 R:Goodner, B.; Hinkley, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
 A.; Lih, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.;
 Science 294, 2323-2338, 2001

A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium
 A:Reference number: A97359; PMID:11743194
 A:Accession: B97483
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-169 <KUR>
 A:Cross-references: GB:AE007869; PIDN:AAK6819.1; PID:g15156027; GSPDB:GN00169
 C:Genetics:
 A:Gene: AGR_C_1859
 A:Map position: circular chromosome

Query Match 100.0%; Score 31; DB 2; Length 169;
 Best Local Similarity 100.0%; Pred. No. 64;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGF 5
 Db 20 GGF 24

RESULT 6

B31227
 nitrogen fixation regulatory protein fixJ - Rhizobium meliloti
 C:Species: Rhizobium meliloti
 C:Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 29-Sep-1999
 C:Accession: B31227; S32838; S3985
 R:David, M.; Davern, M.L.; Batut, J.; Dedieu, A.; Domergue, O.; Ghal, J.; Hertig, C.
 Cell 54, 671-683, 1988
 A:Title: Cascade regulation of nif gene expression in Rhizobium meliloti.
 A:Reference number: A90901; MUID:88311069; PMID:2842062
 A:Accession: B31227
 A:Molecule type: DNA
 A:Residues: 1-204 <DAV>
 A:Cross-references: EMBL:Z21854; NID:g49403; PIDN:CAAT9898.1; PID:g49405
 C:Genetics:
 A:Gene: fixJ
 C:Superfamily: nitrogen fixation regulatory protein fixJ; response regulator homology
 C:Keywords: DNA binding; phosphoprotein; transcription regulation
 F:6-115/Domain: response regulator homology <RR>
 F:54/Binding site: phosphate (Asp) (covalent) #status predicted

Query Match 100.0%; Score 31; DB 2; Length 204;
 Best Local Similarity 100.0%; Pred. No. 78;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGF 5
 Db 199 GGF 203

RESULT 7

E95345
 fixJ Transcriptional activator [imported] - Sinorhizobium meliloti (strain 1021) maga
 C:Species: Sinorhizobium meliloti
 C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
 C:Accession: E95345
 R:Banett, M.V.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; B
 Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
 A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meli
 A:Reference number: A95262; MUID:21396509; PMID:11481432
 A:Accession: E95345
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-204 <KUR>
 A:Cross-references: GB:AE006469; PIDN:AAK65327.1; PID:g14523784; GSPDB:GN00165
 A:Experimental source: strain 1021, megaplasmid pSymA
 R:Galibert, F.; Finan, T.M.; Long, S.R.; Publer, A.; Abola, P.; Ampe, F.; Barloy-Hubl
 l, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.
 Science 293, 668-672, 2001
 A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kles, E.; Komp, C.; Lelau
 heault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh,

A:Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.
A:Reference number: A96039; M0ID:21368234; PMID:11474104
A:Contents: annotation
A:Genetics: fixJ
A:Gene: fixJ
A:Genome: plasmid
C:Superfamily: nitrogen fixation regulatory protein fixJ; response regulator homology

Query Match 100.0%; Score 31; DB 2; Length 204;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGF 5
Db 199 GGF 203

RESULT 8

C86656
glycerol uptake facilitator [imported] - *Lactococcus lactis* subsp. *lactis* (strain IL1403)
C:Species: *Lactococcus lactis* subsp. *lactis*
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 22-Oct-2001
C:Accession: C86656
R:Polotin, A.; Wnacker, P.; Mauger, S.; Jallion, O.; Malarme, K.; Weisenbach, J.; Ehrlich
Genome Res. 11, 731-753, 2001
A:Title: The complete genome sequence of the lactic acid bacterium *Lactococcus lactis* s
A:Reference number: A86625; M0ID:21235186; PMID:11337471
A:Accession: C86656
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-238 <STO>
A:Cross-references: GB:A005176; PID:g12723110; PIDN:AK04349.1; GSPDB:GN00146
A:Experimental source: strain IL1403
C:Genetics: glpP1
A:Gene: glpP1
C:Superfamily: glycerol facilitator protein

Query Match 100.0%; Score 31; DB 2; Length 238;
Best Local Similarity 100.0%; Pred. No. 91;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGF 5
Db 159 GGF 163

RESULT 9

D70660
probable aroE - *Mycobacterium tuberculosis* (strain H37RV)
C:Species: *Mycobacterium tuberculosis*
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: D70660
R:Colet, S.T.; Brosch, R.; Parhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Fellwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998
A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrrell, B.G.
A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome
A:Reference number: A70500; M0ID:98295987; PMID:9634230
A:Accession: D70660
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-269 <COL>
A:Cross-references: GB:Z83863; GB:AL123456; NID:g3261685; PIDN:CA06186.1; PID:g1781050
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: aroE
C:Superfamily: shikimate dehydrogenase; shikimate dehydrogenase homology
F:61-252/Domain: shikimate dehydrogenase homology <SKD>

Query Match 100.0%; Score 31; DB 2; Length 269;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGF 5
Db 53 GGF 57

RESULT 10

C70784
probable dheme cytochrome qrcC - *Mycobacterium tuberculosis* (strain H37RV)
C:Species: *Mycobacterium tuberculosis*
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 16-Jun-2000
C:Accession: C70784
R:Colet, S.T.; Brosch, R.; Parhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Fellwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998
A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrrell, B.G.
A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete geno
A:Reference number: A70500; M0ID:98295987; PMID:9634230
A:Accession: C70784
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-280 <COL>
A:Cross-references: GB:Z70283; GB:AL123456; NID:g3261561; PIDN:CA04263.1; PID:g12370
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: qrcC
C:Superfamily: Streptomyces coelicolor probable dheme cytochrome qrcC; cytochrome c6
C:Keywords: chromoprotein; heme; iron; metalloprotein
F:63-136/Domain: cytochrome c6 homology #status atypical <CYC1>
F:164-235/Domain: cytochrome c6 homology <CYC2>
F:73-76/Binding site: heme (Cys) (covalent) #status predicted
F:77/Binding site: heme iron (His) (axial ligand) #status predicted
F:174-177/Binding site: heme (Cys) (covalent) #status predicted
F:178/Binding site: heme iron (His) (axial ligand) #status predicted

Query Match 100.0%; Score 31; DB 1; Length 280;
Best Local Similarity 100.0%; Pred. No. 11e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGF 5
Db 250 GGF 254

RESULT 11

C87019
hypothetical protein qrcC [imported] - *Mycobacterium leprae*
C:Species: *Mycobacterium leprae*
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 15-Jun-2001
C:Accession: C87019
R:Colet, S.T.; Elgimier, K.; Parhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; R.; Davies, R.M.; Devlin, K.; Dutfoy, S.; Fellwell, T.; Fraser, A.; Hamlin, N.; Holroyd, S.; Nature 409, 1007-1011, 2001
A:Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; A:Title: Massive gene decay in the leprosy bacillus.
A:Reference number: A86903; M0ID:21128732; PMID:11234002
A:Accession: C87019
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-289 <STO>
A:Cross-references: GB:AL450380; NID:g13092954; PIDN:CA031262.1; GSPDB:GN00147
C:Genetics:
A:Gene: qrcC
C:Superfamily: Streptomyces coelicolor probable dheme cytochrome qrcC; cytochrome c6
C:Keywords: chromoprotein; heme; iron; metalloprotein
F:82/85/Binding site: heme (Cys) (covalent) #status predicted
F:86/Binding site: heme iron (His) (axial ligand) #status predicted
F:183-186/Binding site: heme (Cys) (covalent) #status predicted
F:187/Binding site: heme iron (His) (axial ligand) #status predicted

Query Match 100.0%; Score 31; DB 2; Length 289;

Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGFGP 5
|||||
Db 259 GGFGP 263

RESULT 12

A87703
citrate lyase beta subunit, probable [imported] - *Caulobacter crescentus*
C:Species: *Caulobacter crescentus*
C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001
C:Accession: A87703
R:Merlan, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
n, J.; Laub, M.T.; DebRoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gilm, M.L.; Haft, D.H.; Kolon
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of *Caulobacter crescentus*.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: A87703
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-295 <STO>
A:Cross-references: GB:AE005673; NID:g13425417; PIDN:BAF25621.1; GSPDB:GN00148
A:Gene: CC3659
C:Superfamily: *Yersinia pestis* hypothetical 29.7K protein

Query Match 100.0%; Score 31; DB 2; Length 295;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFGP 5
|||||
Db 62 GGFGP 66

RESULT 13

G71088
hypothetical protein PH0972 - *Pyrococcus horikoshii*
C:Species: *Pyrococcus horikoshii*
C>Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 20-Jun-2000
C:Accession: G71088
R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Halkawa, Y.; Hino, Y.; Yamamoto, S.; Seki
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi
DNA Res. 5, 55-76, 1998
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic
A:Reference number: A71000; MUID:98344137; PMID:9679194
A:Accession: G71088
A:Status: Preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-299 <KAP>
A:Cross-references: GB:AP000004; NID:g3236131; PIDN:BAA30069.1; PID:g3257386
A:Experimental source: strain OT3
A>Note: This accession replaces an interim accession for a sequence replaced by GenBank
C:Genetics:
A:Gene: PH0972
C:Superfamily: *Pyrococcus horikoshii* hypothetical protein PH0972

Query Match 100.0%; Score 31; DB 2; Length 299;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFGP 5
|||||
Db 88 GGFGP 92

RESULT 14

F83802
hypothetical protein BH1222 [imported] - *Bacillus halodurans* (strain C-125)
C:Species: *Bacillus halodurans*

C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: F83802
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; H
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* a
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: F83802
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-326 <STO>
A:Cross-references: GB:AP001511; GB:BA000004; NID:g10173727; PIDN:BAF04941.1; GSPDB:G
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH1222

Query Match 100.0%; Score 31; DB 2; Length 326;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFGP 5
|||||
Db 108 GGFGP 112

RESULT 15

D98326
Probable zinc-binding alcohol dehydrogenase (D-mannanase hydrolase) [imported] - *Agro*
C:Species: *Agrobacterium tumefaciens*
C>Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 11-Jan-2002
C:Accession: D98326
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Gold
A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz,
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium*
A:Reference number: A97359; PMID:11743194
A:Accession: D98326
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-348 <KUP>
A:Cross-references: GB:AE007870; PIDN:MAK90134.1; PID:g15160129; GSPDB:GN00170
A:Gene: AGR_L_3122
C:Genetics:
A:Map position: linear chromosome
C:Superfamily: alcohol dehydrogenase; long-chain alcohol dehydrogenase homology

Query Match 100.0%; Score 31; DB 2; Length 348;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFGP 5
|||||
Db 49 GGFGP 53

Search completed: March 17, 2003, 12:11:08
Job time : 15.3333 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 17, 2003, 12:05:19 ; Search time 7 Seconds
(Without alignments)

29.626 Million cell updates/sec

Title: SEQ25-FG

Perfect score: 31

Sequence: 1 g9fgfp 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

112892

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	31	100.0	178	THFL_ARATH	Q9XFB8 arabidopsis
2	31	100.0	204	FIXD_RHIME	P10958 rhizobium m
3	31	100.0	280	QCRC_MYCTU	Q10386 mycobacteri
4	31	100.0	289	QCRC_MYCLE	Q69583 mycobacteri
5	31	100.0	371	KLOM_EISFO	015991 eisenia foe
6	31	100.0	374	PRLB_ACHLY	P27458 achromobact
7	31	100.0	407	YB04_HARIN	P13369 haemophilus
8	31	100.0	430	MAS1_AGRV9	Q28009 agrobacteri
9	31	100.0	512	FUS_BOVIN	P56959 mus taurus
10	31	100.0	518	FUS_MOUSE	P56959 mus musculu
11	31	100.0	526	FUS_HUMAN	P56959 mus musculu
12	31	100.0	638	K22Q_HUMAN	001546 homo sapien
13	31	100.0	645	K22E_HUMAN	P35908 homo sapien
14	31	100.0	723	TR2M_AGRV1	004564 agrobacteri
15	31	100.0	749	TR2M_AGRH4	009109 agrobacteri
16	31	100.0	755	TR2M_AGRV4	P04029 agrobacteri
17	31	100.0	755	TR2N_AGRV1	P25017 agrobacteri
18	31	100.0	764	CPTM_MOUSE	P97742 mus musculu
19	31	100.0	772	CPTM_HUMAN	Q92523 homo sapien
20	31	100.0	772	CPTM_HUMAN	Q92523 homo sapien
21	31	100.0	773	CPT1_HUMAN	Q63704 ratius norv
22	31	100.0	773	CPT1_HUMAN	P50416 homo sapien
23	31	100.0	791	Y046_UREPA	P32198 ratius norv
24	31	100.0	910	PERT_BORPE	099199 ureaplasma
25	31	100.0	911	PERT_BORPE	P14283 bordetella
26	31	100.0	911	PERT_BORBR	003035 bordetella
27	31	100.0	922	PERT_BORBR	P43328 bordetella
28	31	100.0	1093	SM5B_MOUSE	Q60519 mus musculu
29	31	100.0	1228	C1BA_BACTK	P05517 bacillus th
30	31	100.0	4473	PLE1_CRIGR	Q91155 cricetus
31	31	100.0	4684	PLE1_HUMAN	Q15149 homo sapien
32	31	100.0	4687	PLE1_RAT	P04047 ratius norv
33	28	90.3	40	SGP1_THIRO	P56623 thioacapa r
	28	90.3	115	CH15_DROME	P07185 drosophila

ALIGNMENTS

RESULT 1	THFL_ARATH	STANDARD:	PRT:	178 AA.
AC	Q9XFB8;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Thioredoxin F-type 1, chloroplast precursor (TRX-F1).			
OS	Arabidopsis thaliana (Mouse-ear cress).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;			
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsids.			
OX	NCBI_TaxID=3702;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Verdoug L., Meyer Y.;			
RT	Characterisation of the Arabidopsis thaliana thioredoxin f family.;			
RL	Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.			
CC	-1- FUNCTION: Participates in various redox reactions through the			
CC	reversible oxidation of the active center disulfid to a disulfide.			
CC	The F form is known to activate a number of enzymes of the			
CC	photosynthetic carbon cycle (By similarity).			
CC	-1- SUBCELLULAR LOCATION: Chloroplast (By similarity).			
CC	-1- SIMILARITY: BELONGS TO THE THIOREDOXIN FAMILY. PLANT F-TYPE.			
CC	-1- This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.lsb-sib.ch/announce/			
CC	or send an email to license@sib-sib.ch).			
CC	-----			
DR	EMBL; AF144385; AAD35003.1; -			
DR	HSSP; P10599; 1ATU.			
DR	InterPro; IPR000063; ThioRed.			
DR	Pfam; PF00085; ThioRed; 1.			
DR	PRINTS; PR00421; THIOREDOXIN.			
DR	PROSITE; PS00194; THIOREDOXIN; 1.			
KW	Redox-active center; Electron transport; Chloroplast; Transit peptide;			
KW	Multigene family.			
FT	TRANSIT 1			
FT	CHAIN 1			
FT	DISULFID 99			
FT	SEQUENCE 178 AA; 19351 MW; 7B4ECFE9B6200F8 CRC64;			
QY	Query Match	100.0%;	Score 31; DB 1; Length 178;	
QY	Best Local Similarity	100.0%;	Pred. No. 39;	
QY	Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
DB	19 GGFSP 23			

RESULT 2

FIXJ_RHIME STANDARD: PRT; 204 AA.
 AC P10958;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Transcriptional regulatory protein fixJ.
 GN FIXJ OR RA0669 OR SMA1227.
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC Plasmid pSymba (megaplasmid 1).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Sinorhizobium.
 OX NCBI_TaxID=382;
 RN
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88311069; PubMed=2842062;
 RA David M., Daveran M.-L., Batur J., Dedieu A., Domergue O., Ghat J.,
 RA Hertig C., Bolstead P., Kahn D.;
 RT "Cascade regulation of nif gene expression in Rhizobium meliloti.";
 RL Cell 54:671-683(1988).
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=1021;
 RX MEDLINE=21396509; PubMed=11481432;
 RA Barnett M.J., Fisher R.F., Jones T., Komp C., Abola A.P.,
 RA Barloy-Hubler F., Bowser L., Capella D., Gallibert F., Gouzy J.,
 RA Gurjal M., Hong A., Huzar L., Hyman R.W., Kahn D., Kahn M.L.,
 RA Kalman S., Keating D.H., Palm C., Peck M.C., Surzycki R., Wells D.H.,
 RA Yen K.-C., Davis R.W., Federpspiel N.A., Long S.R.;
 RT "Nucleotide sequence and predicted functions of the entire
 RT Sinorhizobium meliloti pSymba megaplasmid.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9883-9888(2001).
 RN
 RP SEQUENCE OF 200-204 FROM N.A.
 RX MEDLINE=89305532; PubMed=2663474;
 RA Batur J., Daveran-Mingot M.-L., David M., Jacobs J., Garnerone A.-M.,
 RA Kahn D.;
 RT "fixK, a gene homologous with fur and cfp from Escherichia coli,
 RT regulates nitrogen fixation genes both positively and negatively in
 RL Rhizobium meliloti.";
 RN EMBO J. 8:1279-1286(1989).
 RN
 RP REVIEW, AND MUTAGENESIS OF HTH REGION.
 RX MEDLINE=91312138; PubMed=1857213;
 RA Kahn D., Ditta G.S.;
 RT "Modular structure of fixJ: domain of the transcriptional activator
 RT Mol. Microbiol. 5:987-997(1991).
 RN
 RP X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS) OF 1-126.
 RX MEDLINE=20113479; PubMed=10647182;
 RA Gouet P., Fabry B., Guillet V., Black C., Mourey L., Kahn D.,
 RA Samana J.-P.;
 RT "Structural transitions in the fixJ receiver domain.";
 RL Structure 7:1517-1526(1999).
 RN
 RP FUNCTION: FIXJ, WHEN ACTIVATED BY FIXL, INDUCES THE EXPRESSION OF
 CC BOTH NIFA, REQUIRED FOR ACTIVATION OF CLASSICAL NIF AND FIX GENES,
 CC AND FIXK, REQUIRED FOR FIXN ACTIVATION.
 CC
 CC SUBCELLULAR LOCATION: Cytoplasmic (Probable).
 CC
 CC PTM: PHOSPHORYLATED BY FIXL.
 CC
 CC SIMILARITY: CONTAINS 1 RESPONSE REGULATORY DOMAIN.
 CC
 CC SIMILARITY: BELONGS TO THE LUXR/UHPA FAMILY OF TRANSCRIPTIONAL
 CC REGULATORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
 CC or send an email to license@sib-sib.ch)
 CC
 CC EMBL; J03174; AAA26276.1; -

DR EMBL; Z21854; CAAY9898.1; -
 DR EMBL; AE007255; AAK65327.1; -
 DR EMBL; X15079; CAA33182.1; -
 DR PIR; B31227; B31227.
 DR PIR; S39985; S39985.
 DR PIR; S32838; S32838.
 DR PDB; 1DBW; 26-NOV-99.
 DR PDB; 1DCK; 26-NOV-99.
 DR InterPro; IPR000792; HTH_LuxR.
 DR InterPro; IPR001789; Response_reg.
 DR Pfam; PF00072; response_reg.1.
 DR Pfam; PF00196; GcR; 1.
 DR PRINTS; PR00038; HTHLUXR.
 DR ProDom; PD000039; Response_reg.1.
 DR ProDom; PD000307; HTH_LuxR; 1.
 DR SMART; SM00421; HTH_LuxR; 1.
 DR SMART; SM00448; REC; 1.
 DR PROSITE; PS00622; HTH_LuxR_FAMILY.1.
 DR PROSITE; PS50110; RESPONSE_REGULATORY.1.
 DR Sensor; transduction; Phosphorylation; Transcription regulation;
 KW DNA-binding; Activator; Nitrogen fixation; Plasmid; 3D-structure;
 KW Complete proteome.
 FT DOMAIN 1 119 RESPONSE REGULATORY.
 FT MOD_RES 54 54 PHOSPHORYLATION.
 FT DNA_BIND 159 178 H-T-H MOTIF (BT SIMILARITY).
 SQ SEQUENCE 204 AA; 22218 MW; 2EDA356967352292 CRC64;
 Query Match 100.0%; Score 31; DB 1; Length 204;
 Best local Similarity 100.0%; Pred. No. 44;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 GGRGP 5
 Db 199 GGRGP 203
 Db 199 GGRGP 203
 RESULT 3
 ID QCRC_MYCTU STANDARD: PRT; 280 AA.
 AC 010386;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Ubiquinol-cytochrome c reductase cytochrome c subunit.
 GN QCRC OR RV2194 OR MT2250 OR MTCY190.05.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37Rv;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
 RA Baddock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltham T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence.";
 RL Nature 393:537-544(1998).
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / OshKosh;
 RA Fieisemann R.D., Ashdosh;
 RA Peterson J.F., Deboy R., Dodson R., Gwin M.L., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
 RA Belcher A., Uterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and

```

RT laboratory strains."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC
CC -1- SUBUNIT: THE MAIN SUBUNITS OF THE UBIQUINOL-CYTOCHROME C COMPLEX
CC ARE: CYTOCHROME B, THE RIESKE PROTEIN AND A CYTOCHROME C SUBUNIT.
CC
CC -1- SIMILARITY: SOME, TO CYTOCHROME C.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL: Z70283; CAA94263.1; -.
DR EMBL: AB007071; AAK4636.1; -.
DR TIGR: MT2250; -.
DR Tuberculin; RV2194; -.
DR InterPro: IPR000345; CytC_heme_bind.
DR InterPro: IPR003088; Cyt.CI.
DR Pfam: PF00034; cytochrome_c; 1.
DR ProDom: PD004020; Cyt.CIC; 1.
DR PROSITE: PS00190; CYTOCHROME_C; 2.
KW Electron transport; Heme; Transmembrane; Complete proteome.
FT BINDING 73 73 HEME 1 (COVALENT) (POTENTIAL).
FT BINDING 76 76 HEME 1 (COVALENT) (POTENTIAL).
FT METAL 77 77 IRON 1 (HEME AXIAL LIGAND) (POTENTIAL).
FT BINDING 174 174 HEME 2 (COVALENT) (POTENTIAL).
FT BINDING 177 177 HEME 2 (COVALENT) (POTENTIAL).
FT METAL 178 178 IRON 2 (HEME AXIAL LIGAND) (POTENTIAL).
FT TRANSMEM 25 45 POTENTIAL.
FT TRANSMEM 258 278 POTENTIAL.
SQ SEQUENCE 280 AA; 29138 MW; 7A058AEFE12EF713 CRC64;

Query Match 100.0%; Score 31; DB 1; Length 280;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFGP 5
Db 250 GGFGP 254

RESULT 4
QCRC_MYCLE STANDARD; PRT; 289 AA.
AC 069583;
DT 30-MAY-2000 (Rel. 39, Created)
DT 15-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ubiquinol-cytochrome c reductase cytochrome c subunit.
OS QCRC OR ML0881 OR MLC8268.36.
GN Mycobacterium lepreae.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID:1769;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-TN;
RX MEDLINE-2118732; PubMed-11234002;
RA Cole S.T., Elgmeler K., Parthill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrall B.G.;
RA "Massive gene decay in the leprosy bacillus.";
RT Nature 409:1007-1011(2001).
CC -1- SUBUNIT: THE MAIN SUBUNITS OF THE UBIQUINOL-CYTOCHROME C COMPLEX
CC ARE: CYTOCHROME B, THE RIESKE PROTEIN AND A CYTOCHROME C SUBUNIT.

```

```

CC
CC -1- SIMILARITY: SOME, TO CYTOCHROME C.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL: AL583920; CAC31262.1; -.
DR EMBL: AL022602; CAA18702.1; -.
DR HSSP: P56534; 1c68.
DR Lepronia; ML0881; -.
DR InterPro: IPR000345; CytC_heme_bind.
DR InterPro: IPR003088; Cyt.CI.
DR Pfam: PF00034; cytochrome_c; 1.
DR ProDom: PD004020; Cyt.CIC; 1.
DR PROSITE: PS00190; CYTOCHROME_C; 2.
KW Electron transport; Heme; Transmembrane; Complete proteome.
FT BINDING 82 82 HEME 1 (COVALENT) (POTENTIAL).
FT BINDING 85 85 HEME 1 (COVALENT) (POTENTIAL).
FT METAL 86 86 IRON 1 (HEME AXIAL LIGAND) (POTENTIAL).
FT BINDING 183 183 HEME 2 (COVALENT) (POTENTIAL).
FT BINDING 186 186 HEME 2 (COVALENT) (POTENTIAL).
FT METAL 187 187 IRON 2 (HEME AXIAL LIGAND) (POTENTIAL).
FT TRANSMEM 37 55 POTENTIAL.
FT TRANSMEM 267 287 POTENTIAL.
SQ SEQUENCE 289 AA; 30672 MW; 224D52A3A40B5314 CRC64;

Query Match 100.0%; Score 31; DB 1; Length 289;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFGP 5
Db 259 GGFGP 263

RESULT 5
KLOM_EISFO STANDARD; PRT; 371 AA.
AC 015991;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Lombriclike kinase (EC 2.7.3.5) (LK).
OS Eisenia foetida (Common brandling worm) (Common dung-worm).
OC Eukaryota; Metazoa; Annelida; Clitellata; Oligochaeta; Haplotaxida;
OC Lumbricidae; Lumbricolidae; Eisenia.
OX NCBI_TaxID:6396;
RN [1]
RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
RX MEDLINE-98094195; PubMed-9434106;
RA Suzuki T., Kawasaki Y., Furukohri T., Ellington W.R.;
RA "Evolution of phosphagen kinase. VI. Isolation, characterization and
RT cDNA-derived amino acid sequence of lombriclike kinase from the
RT earthworm Eisenia foetida, and identification of a possible candidate
RT for the guanidine substrate recognition site.";
RL Biochem. Biophys. Acta 1343:152-159(1997).
CC -1- CATALYTIC ACTIVITY: ATP + Lombriclike.
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE ATP-GUANIDO PHOSPHOTRANSFERASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL: AB008013; BAA22872.1; -.

```

DR HSP: P00563; 2CRK.
 DR InterPro: IPR000749; ATP-gua_Ptrans.
 DR Pfam: PF00217; ATP-gua_Ptrans. 1.
 DR Pfam: PF02807; ATP-gua_Ptrans. 1.
 DR PROSITE: PS00112; GUANIDO_KINASE. 1.
 KW Kinase: Transferase.
 FT ACT_SITE 271
 SO SEQUENCE 371 AA; 41781 MW; 8F8409D72F92ACB9 CRC64;

Query Match 100.0%; Score 31; DB 1; Length 371;
 Best Local Similarity 100.0%; Pred. No. 75;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFSP 5
 DB 86 GGFSP 90

RESULT 6
 ID PRIB_ACHLY STANDARD; PRT; 374 AA.
 AC P27438;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Beta-lytic metalloendopeptidase precursor (EC 3.4.24.32) (Beta-lytic
 DE protease).
 OS Achromobacter lyticus.
 OC Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;
 OC Achromobacter.
 OX NCBI_TaxID=224;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 196-220.
 RC STRAIN=MA97-1;
 RX MEDLINE=91035265; PubMed=2228973;
 RA Li S.L., Norioka S., Sakiyama F.;
 RT "Molecular cloning and nucleotide sequence of the beta-lytic protease
 RT gene from Achromobacter lyticus.";
 RL J. Bacteriol. 172:6506-6511(1990).
 CC -1- CATALYTIC ACTIVITY: Cleavage of N-acetyl-L-muamoyl-L-Ala, and of the
 CC Insulin B chain at 23-Gly-L-Phe-24 > 18-Val-L-Cys(SO(3)H).
 CC -1- CORRELATION: BINDS ONE ZINC ION.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M23.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch)
 CC -----
 CC EMBL: M60896; AAA21906.1; -
 DR PIR: A37151; LYXYLY.
 DR MEROPS: M23.001; -
 DR InterPro: IPR000841; Blytic_Mendep.
 DR PRINTS: PR00933; BLYTICPTASE.
 KW Hydrolase; Metalloprotease; Zinc; Zymogen; Signal.
 FT SIGNAL 1 24
 FT PROPEP 25 195
 FT CHAIN 196 374
 FT DISULFID 261 307 BETA-LYTIC METALLOENDOPEPTIDASE.
 FT DISULFID 351 364 BY SIMILARITY.
 FT METAL 316 316 ZINC (POTENTIAL).
 FT METAL 318 318 ZINC (POTENTIAL).
 SQ SEQUENCE 374 AA; 40084 MW; 431E51B84575DE14 CRC64;

OY 1 GGFSP 5

DB 180 GGFSP 184
 RESULT 7
 ID YB04_HAEIN STANDARD; PRT; 407 AA.
 AC P71369;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical metalloprotein Hili04.
 CN Hili04.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Haemophilus.
 OX NCBI_TaxID=727;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Rd / KW20 / ATCC 51907;
 RX MEDLINE=95350630; PubMed=7542800;
 RA Frieschmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Goessens J.D.,
 RA Scott J.D., Shirley R., Liu L.-I., Glodex A., Kelley J.M.,
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RA Utterback T.R., Hanna M.C., Nguyen D.T., Sauder D.M., Brandon R.C.,
 RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghegan N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.;
 RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
 RT Rd.";
 RL Science 269:496-512(1995).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
 CC (Potential).
 CC -----
 CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch)
 CC -----
 CC EMBL: U32790; AAC22759.1; -
 DR TIGR: H11104.
 DR InterPro: IPR004746; Bnate-transport.
 DR Pfam: PF00083; sugar_tr. 1.
 DR TIGRFAMS: TIGR00895; 2A0115; 1.
 DR PROSITE: PS00216; SUGAR_TRANSPORT_1; FALSE_NEG.
 DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
 KW Hypothetical protein; Transport; Transmembrane; Inner membrane;
 KW Complete proteome.
 FT DOMAIN 1 16 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 17 37 1 (POTENTIAL).
 FT DOMAIN 38 48 PERIPLASMIC (POTENTIAL).
 FT TRANSSEM 49 69 2 (POTENTIAL).
 FT DOMAIN 70 77 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 78 98 3 (POTENTIAL).
 FT DOMAIN 99 107 PERIPLASMIC (POTENTIAL).
 FT TRANSSEM 108 128 4 (POTENTIAL).
 FT DOMAIN 129 138 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 139 159 5 (POTENTIAL).
 FT DOMAIN 160 160 PERIPLASMIC (POTENTIAL).
 FT TRANSSEM 161 181 6 (POTENTIAL).
 FT DOMAIN 182 224 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 225 245 7 (POTENTIAL).
 FT DOMAIN 246 261 PERIPLASMIC (POTENTIAL).
 FT TRANSSEM 262 282 8 (POTENTIAL).
 FT DOMAIN 283 288 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 289 309 9 (POTENTIAL).

```

FT DOMAIN 310 312 PERIPLASMIC (POTENTIAL).
FT TRANSSEM 333 333 10 (POTENTIAL).
FT TRANSSEM 334 357 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 358 378 11 (POTENTIAL).
FT TRANSSEM 399 399 12 (POTENTIAL).
FT DOMAIN 400 407 CYTOPLASMIC (POTENTIAL).
SQ SEQUENCE 407 AA; 43681 MW; 618586C6B37CA49 CRC64;

Query Match
Best Local Similarity 100.0%; Score 31; DB 1; Length 407;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFSP 5
Db 360 GGFSP 364

RESULT 8
MASI_AGR9 STANDARD; PRT: 430 AA.
ID MASI_AGR9
AC P50202;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Agropine synthesis reductase (EC 1.-.-.-).
GN MASI.
OS Agrobacterium tumefaciens (strain 15955).
OC plasmid pTI15955.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
ON NCBI_TaxID=190386;
RX [1]
RP SEQUENCE FROM N.A.
RA Barker R.F., Idler K.B., Thompson D.V., Kemp J.D.;
RT "Nucleotide sequence of the T-DNA region from the Agrobacterium
RT tumefaciens octopline T1 plasmid pTI15955."
RL Plant Mol. Biol. 2:335-350(1983).
CC -1- FUNCTION: REDUCES DEOXY-FRUCTOSYL-GLUTAMINE TO MANNOPINE.
CC -1- PATHWAY: AGROPINE / MANNOPINE SYNTHESIS.
CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
CC (SDR) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X00493; CAA25187.1; -
DR InterPro: IPR002198; ADH_short.
DR InterPro: IPR001345; PG/BPGM_mutase.
DR Pfam: PF00106; adh_short; 1.
DR Pfam: PF00300; PGAM; 1.
DR PROSITE: PS00061; ADH_SHORT; 1.
DR KMW Plasmid: Oxidoreductase.
FT NF_BIND 203 227 NAD OR NADP (BY SIMILARITY).
FT ACT SITE 346 346 BY SIMILARITY.
SQ SEQUENCE 430 AA; 46750 MW; CE47F536BD92AAD CRC64;

Query Match
Best Local Similarity 100.0%; Score 31; DB 1; Length 430;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFSP 5
Db 115 GGFSP 119

RESULT 9
FUS_BOVIN STANDARD; PRT: 512 AA.
ID FUS_BOVIN

```

```

AC 028009;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE RNA-binding protein FUS (P1gpen protein).
GN FUS.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
ON NCBI_TaxID=9913;
RX [1]
RP SEQUENCE FROM N.A.
RA TISSUE-Aorta;
RX MEDLINE-96175600; PubMed-8631501;
RA Allegro M.C., Allegro M.A.;
RT "A nuclear protein regulated during the transition from active to
RT quiescent phenotype in cultured endothelial cells."
RL Dev. Biol. 174:288-297(1996).
CC -1- FUNCTION: BINDS BOTH SINGLE-STRANDED AND DOUBLE-STRANDED DNA AND
CC PROMOTES ATP-INDEPENDENT ANNEALING OF COMPLEMENTARY SINGLE-
CC STRANDED DNAs AND D-LOOP FORMATION IN SUPRAPHILICAL DOUBLE-STRANDED
CC DNA. MAY PLAY A ROLE IN MAINTENANCE OF GENOMIC INTEGRITY (BY
CC SIMILARITY).
CC -1- SUBUNIT: COMPONENT OF NUCLEAR RIBOPROTEIN COMPLEXES (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: NUCLEAR, EXHIBITS DIFFUSE STAINING
CC THROUGHOUT (EXCLUDING NUCLEOLI), TOGETHER WITH A SMALL NUMBER OF
CC INTENSELY STAINED FOCAL POINTS, OR GRANULES, AND PUNCTATE STAINING
CC ALONG THE NUCLEAR ENVELOPE.
CC -1- DOMAIN: THE C-TERMINAL DOMAIN BINDS CARBOHYDRATES.
CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
CC -1- SIMILARITY: CONTAINS 1 RANBP2-TYPE ZINC FINGER.
CC -1- SIMILARITY: BELONGS TO THE TET FAMILY OF RNP PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U26024; AAC13543.1; -
DR HSP: P09651; 1HA1.
DR InterPro: IPR000504; RNA_rec_mot.
DR InterPro: IPR001876; znf_RanGDP.
DR Pfam: PF00076; rrm; 1.
DR Pfam: PF00641; zf-RanBP; 1.
DR SMART: SM00360; RRM; 1.
DR SMART: SM00547; znf_RBZ; 1.
DR PROSITE: PS50102; RRM; 1.
DR PROSITE: PS00030; RRM_RNP_1; FALSE_NEG.
DR PROSITE: PS01358; zf_RANBP2_1; 1.
DR PROSITE: PS50199; zf_RANBP2_2; 1.
DR RNA-binding: DNA-binding; Nuclear protein; Repeat; zinc-finger; zinc;
KW Metal-binding.
FT DOMAIN 1 164 GLN/GLY/SER/TYR-RICH.
FT DOMAIN 165 253 GLY-RICH.
FT DOMAIN 271 357 RNA-BINDING (RRM).
FT DOMAIN 357 512 ARG/GLY-RICH.
FT ZN_FING 408 439 RANBP2-TYPE.
SQ SEQUENCE 512 AA; 52240 MW; 3652329C044F1386 CRC64;

Query Match
Best Local Similarity 100.0%; Score 31; DB 1; Length 512;
Matches 5; Conservative 0; Mismatches 1e+02;

```

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGFSP 5
|||||
Db 490 GGFSP 494

RESULT 10
ID FUS_MOUSE STANDARD; PRT: 518 AA.
AC P56959;
DT 16-OCT-2001 (Rel. 40, Created)
DT 15-OCT-2001 (Rel. 40, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE RNA-binding protein FUS (P19gen protein).
GN FUS.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN (1)
RP SEQUENCE FROM N.A.
RA Alapat S.R., Zhang M., Zhao X., Allegro M.A., Allegro M.C.,
RA Burdall C.A.;
RT *Regulation of p19gen expression in mouse embryos.*;
RT Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: BINDS BOTH SINGLE-STRANDED AND DOUBLE-STRANDED DNA AND
CC PROMOTES ATP-INDEPENDENT ANNEALING OF COMPLEMENTARY SINGLE-
CC STRANDED DNAs AND D-LOOP FORMATION IN SUPERHELICAL DOUBLE-STRANDED
CC DNA. MAY PLAY A ROLE IN MAINTENANCE OF GENOMIC INTEGRITY (BY
CC SIMILARITY).
CC -1- SUBUNIT: COMPONENT OF NUCLEAR RIBOPROTEIN COMPLEXES. INTERACTS
CC WITH IIF3 (By similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
CC -1- SIMILARITY: CONTAINS 1 RANBP2-TYPE ZINC FINGER.
CC -1- SIMILARITY: BELONGS TO THE TER FAMILY OF RNP PROTEINS.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib.ch).
CC -----
DR EMBL: AF224264; AAF70602.1; .
DR HSP: P09651; IHA1.
DR MGI: MGI:1353633; FUS.
DR InterPro: IPR000504; RNA_rec_mot.
DR InterPro: IPR01876; znf_RangBP.
DR Pfam: PF00076; rrm; 1. znf_RangBP.
DR SMART: PF00641; zF-RanBP; 1.
DR SMART: SM00360; RRM; 1.
DR SMART: SM00547; znf_RBZ; 1.
DR PROSITE: PS0102; RRM; 1.
DR PROSITE: PS00030; RRM_RNP_1; FALSE_NEG.
DR PROSITE: PS01359; zF_RANBP2_1; 1.
DR PROSITE: PS01359; zF_RANBP2_2; 1.
DR RNA-binding; DNA-binding; Nuclear protein; Repeat; Zinc-finger; Zinc;
KM Metal-binding.
KM Domain 1 167 GLN/GLY/SER/TYR-RICH.
FT DOMAIN 168 265 GLY-RICH.
FT DOMAIN 278 364 RNA-BINDING (RRM).
FT DOMAIN 364 518 ARG/GLY-RICH.
FT ZN_FING 415 446 RANBP2-TYPE.
SQ SEQUENCE 518 AA: 52673 MW: E06F231BEEED78D6 CRC64;

Query Match 100.0%; Score 31; DB 1; Length 518;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGFSP 5

Db 496 GGFSP 500
|||||

RESULT 11
ID FUS_HUMAN STANDARD; PRT: 526 AA.
AC P35637;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE RNA-binding protein FUS (Oncogene FUS) (Oncogene TRS) (Translocated in
DE liposarcoma protein) (POMP75) (75 kDa DNA-pairing protein).
GN FUS OR TLS.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN (1)
RP SEQUENCE FROM N.A. (LONG ISOFORM).
RX MEDLINE-93288139; PubMed-8510758;
RA Crozat A., Aman P., Mandahl N., Ron D.;
RT *Fusion of CHOP to a novel RNA-binding protein in human myxoid
RT liposarcoma.*;
RT Nature 363:640-644(1993).
RN (2)
RP SEQUENCE FROM N.A. (SHORT ISOFORM).
RX MEDLINE-93350637; PubMed-7503811;
RA Rabbits T.H., Forster A., Larson R., Nathan P.;
RT *Fusion of the dominant negative transcription regulator CHOP with a
RT novel gene FUS by translocation t(12;16) in malignant liposarcoma.*;
RT Nat. Genet. 4:175-180(1993).
RN (3)
RP SEQUENCE FROM N.A. (LONG AND SHORT ISOFORMS).
RX MEDLINE-99013873; PubMed-9795213;
RA Morinoshiri F., Ootsuka Y., Arai K., Ichikawa H., Mitani S.,
RA Munakata N., Ohki M.;
RT *Genomic structure of the human RBP56/htraf168 and FUS/TLS genes.*;
RT Gene 221:191-198(1998).
RN (4)
RP SEQUENCE OF 235-244; 307-312; 335-345 AND 349-357, AND
RP CHARACTERIZATION.
RX MEDLINE-20036580; PubMed-10567410;
RA Baehndorf H., Knudsen M., Sok J., Ron D., Lopez B.S., Akimov A.T.;
RT *Human 75-kDa DNA-pairing protein is identical to the pro-oncoprotein
RT TLS/FUS and is able to promote D-loop formation.*;
RT J. Biol. Chem. 274:34337-34342(1999).
RN (5)
RP SEQUENCE OF 265-276; 317-331 AND 335-357, AND IDENTIFICATION.
RX MEDLINE-99369251; PubMed-10442642;
RA Bertrand P., Akimov A.T., Delacote F., Durbach A., Lopez B.S.;
RT *Human POMP75 is identified as the pro-oncogene TRS/FUS: both
RT POMP75 and POMP100 DNA homologous pairing activities are associated
RT to cell proliferation.*;
RT Oncogene 18:4515-4521(1999).
RN (6)
RP CHROMOSOMAL TRANSLOCATION.
RX MEDLINE-94243799; PubMed-8187069;
RA Ichikawa H., Shimizu K., Hayashi Y., Ohki M.;
RT *An RNA-binding protein gene, TLS/FUS, is fused to ERG in human
RT myeloid leukemia with t(16;21) chromosomal translocation.*;
RT Cancer Res. 54:2865-2868(1994).
RN (7)
RP -1- FUNCTION: BINDS BOTH SINGLE-STRANDED AND DOUBLE-STRANDED DNA AND
RP PROMOTES ATP-INDEPENDENT ANNEALING OF COMPLEMENTARY SINGLE-
RP STRANDED DNAs AND D-LOOP FORMATION IN SUPERHELICAL DOUBLE-STRANDED
RP DNA. MAY PLAY A ROLE IN MAINTENANCE OF GENOMIC INTEGRITY.
RP -1- SUBUNIT: COMPONENT OF NUCLEAR RIBOPROTEIN COMPLEXES. INTERACTS
RP WITH IIF3.
RP -1- SUBCELLULAR LOCATION: Nuclear.
RP -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
RP SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
RP -1- TISSUE SPECIFICITY: UNOBTAINED.
RP -1- DISEASE: A FORM OF MALIGNANT MYXOID LIPOSARCOMA IS CHARACTERIZED

BY A CHROMOSOMAL TRANSLOCATION T(12;16)(Q13;P11) THAT INVOLVES FUS AND CHOP.

-1- DISEASE: A FORM OF ACUTE MYELOID LEUKEMIA (AML) IS CHARACTERIZED BY A CHROMOSOMAL TRANSLOCATION T(16;21)(P11;Q22) THAT INVOLVES FUS AND ERG.

-1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).

-1- SIMILARITY: CONTAINS 1 RANBP2-TYPE ZINC FINGER.

-1- SIMILARITY: BELONGS TO THE TET FAMILY OF RNP PROTEINS.

THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; S62140; AAB27102.1; .

EMBL; S62138; AAB27103.1; ALT_SEQ.

EMBL; X71427; CAA50558.1; ALT_SEQ.

EMBL; X71428; CAA50559.1; ALT_SEQ.

EMBL; AF071213; AAC35285.1; .

EMBL; AF071213; AAC35284.1; .

PIR; S33799; S33799.

HSP; P09651; 1HAI.

Genew; HGNC:4010; FUS.

MIM; 137070; .

MIM; 151900; .

InterPro: IPR000504; RNA_rec_mot.

InterPro: IPR001876; Znf_RanGDP.

Pfam; PF00076; trm; 1.

Pfam; PF00641; zf-RanBP; 1.

SMART; SM00360; RRM; 1.

SMART; SM00547; ZNF_RBZ; 1.

PROSITE; PS0102; RRM; 1.

PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.

PROSITE; PS01358; ZF_RANBP2_1; 1.

PROSITE; PS01358; ZF_RANBP2_2; 1.

Proto-oncogene; RNA-binding; DNA-binding; Nuclear protein; Repeat; Metal-binding; Chromosomal translocation; Zinc-finger; Zinc;

DOMAIN 1 165 GLN/GLY/SER/TYR-RICH.

DOMAIN 166 267 GLY-RICH.

DOMAIN 285 371 RNA-BINDING (RRM).

DOMAIN 371 526 ARG/GLY-RICH.

ZNFING 422 453 RANBP2-TYPE.

SITE 266 267 BREAKPOINT FOR TRANSLOCATION TO FORM FUS/TLS-CHOP ONCOGENE.

VARSPLIC 64 65 TG -> S (IN SHORT ISOFORM).

CONFLICT 338 338 T -> N (IN REF. 4).

SEQUENCE 526 AA; 53426 MW; 53482 CB; 905549 CRC64;

Query Match 100.0%; Score 31; DB 1; Length 526;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFEP 5
|||||

Db 504 GGFEP 508

RESULT 12
K220_HUMAN STANDARD; PRT; 638 AA.

AC 001546;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Keratin, type II cytoskeletal 2 oral (Cytokeratin 2p) (K2p)
DE (CK 2p).
GN KRT2P.
OS Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Primates; Catarrhini; Homiidae; Homo.
NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
RX MEDLINE-93114504; Pubmed-1282112;
RA Collin C., Ouhayoun J.P., Grund C., Franke W.W.;
RT "Suprabasal marker proteins distinguishing keratinizing squamous
epithelia: cytokeratin 2 polypeptides of oral masticatory epithelium
and epidermis are different."
RT Differentiation 51:137-148(1992).
CC -1- FUNCTION: PROBABLY CONTRIBUTES TO TERMINAL CORNIFICATION.
CC -1- SUBUNIT: HETEROHEXAMER OF TWO TYPE I AND TWO TYPE II KERATINS.
CC -1- DEVELOPMENTAL STAGE: SYNTHESIZED DURING MATURATION OF EPIDERMAL
KERATINOCYTES.
CC -1- MISCELLANEOUS: THERE ARE TWO TYPES OF CYTOSKELETAL AND
MICROFIBRILLAR KERATIN: I (ACIDIC; 40-55 kDa) [K9 TO K20] AND II
(NEUTRAL TO BASIC; 56-70 kDa) [K1 TO K8].
CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.

THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; M99063; AAA35746.1; .

InterPro: IPR001644; IF.

InterPro: IPR003054; Keratin_II.

Pfam; PF00038; filament; 1.

PRINTS; PR01276; TYPE2KERATIN.

PROSITE; PS00226; IF; 1.

KW Intermediate filament; Coiled coil; Keratin; Phosphorylation.

FT DOMAIN 1 182 HEAD.

FT DOMAIN 183 492 ROD.

FT DOMAIN 493 638 TAIL.

FT DOMAIN 183 218 COIL_1A.

FT DOMAIN 219 237 LINKER_1.

FT DOMAIN 238 329 COIL_1B.

FT DOMAIN 330 353 LINKER_12.

FT DOMAIN 354 492 COIL_2.

FT MOD_RES 57 57 PHOSPHORYLATION (BY SIMILARITY).

SEQUENCE 638 AA; 65871 MW; 98743AB8872076AF CRC64;

Query Match 100.0%; Score 31; DB 1; Length 638;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFEP 5
|||||

Db 144 GGFEP 148

RESULT 13
K22E_HUMAN STANDARD; PRT; 645 AA.

AC P35908;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Keratin, type II cytoskeletal 2 epidermal (Cytokeratin 2e) (K2e)
DE (CK 2e).
GN KRT2A OR KRT2E.
OS Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE: Thick epidermis;
RX MEDLINE-92380238; Pubmed-1380918;
RA Collin C., Moll R., Kubicka S., Ouhayoun J.-P., Franke W.W.;

RT "Characterization of human cyokeratin 2, an epidermal cytoskeletal
 RT protein synthesized late during differentiation.",
 RN Exp. Cell Res. 202:132-141(1992).
 RN [2]
 RP SEQUENCE FROM N.A., AND VARIANTS IBS TYR-192 AND LYS-482.
 RX MEDLINE-99019489; PubMed-9804344.
 RA Smith F.J.D., Maling C., Covello S.P., Higgins C., Schmidt M.,
 RA Lane E.B., Utto J., Leigh I.M., McLean W.H.I.;
 RT "Genomic organization and fine mapping of the keratin 2e gene (KRT2E):
 RT K2e VI domain polymorphism and novel mutations in ichthyosis bullosa
 RT of Siemens.",
 RL J. Invest. Dermatol. 111:817-821(1998).
 RN [3]
 RP VARIANT IBS LYS-493.
 RX MEDLINE-94358459; PubMed-7521371;
 RA McLean W.H.I., Morley S.M., Lane E.B., Eady R.A.J., Griffiths W.A.D.,
 RA Paige D.G., Harper J.I., Higgins C., Leigh I.M.;
 RT "Ichthyosis bullosa of Siemens -- a disease involving keratin 2e.",
 RL J. Invest. Dermatol. 103:277-281(1994).
 RN [4]
 RP VARIANTS IBS PRO-187, PRO-490 AND LYS-493.
 RX MEDLINE-94358461; PubMed-8077693;
 RA Kremer H., Zeeuwen P., McLean W.H.I., Mariman E.C.M., Lane E.B.,
 RA van de Kerkhof P.C.M., Rogers H.-H., Steijlen P.M.;
 RT "Ichthyosis bullosa of Siemens is caused by mutations in the keratin
 RT 2e gene.",
 RL J. Invest. Dermatol. 103:286-289(1994).
 RN [5]
 RP VARIANTS IBS ASP-493 AND LYS-493.
 RX MEDLINE-95038833; PubMed-7524919;
 RA Rothnagel J.A., Traupe H., Wojcik S., Huber M., Hohl D.,
 RA Pitterkow M.R., Saeki H., Ishibashi Y., Roop D.R.;
 RT "Mutations in the rod domain of keratin 2e in patients with ichthyosis
 RT bullosa of Siemens.",
 RL Nat. Genet. 7:485-490(1994).
 RN [6]
 RP VARIANT IBS PRO-485.
 RX MEDLINE-97348962; PubMed-9204966;
 RA Yang J.-M., Lee S., Bang H.-D., Kim W.-S., Lee E.-S., Steinhert P.M.;
 RT "A novel threonine-to-proline mutation at the end of 2B rod domain in
 RT the keratin 2e chain in ichthyosis bullosa of Siemens.",
 RL J. Invest. Dermatol. 109:116-118(1997).
 RN [7]
 RP VARIANT IBS ASN-188.
 RX MEDLINE-99181809; PubMed-10084318;
 RA Afifi M.J., Longley M.A., Epstein E.H. Jr., Scott G., Goldsmith L.A.,
 RA Rothnagel J.A., Roop D.R.;
 RT "A novel mutation in the 1A domain of keratin 2e in ichthyosis bullosa
 RT of Siemens.",
 RL J. Invest. Dermatol. 112:380-382(1999).
 RN [8]
 RP VARIANT IBS ASP-192.
 RX MEDLINE-20087079; PubMed-10620137;
 RA Takizawa Y., Akiyama M., Nagashima M., Shimizu H.;
 RT "A novel aspartic acid mutation in the rod 1A domain in
 RT keratin 2e in a Japanese family with ichthyosis bullosa of Siemens.",
 RL J. Invest. Dermatol. 114:193-195(2000).
 CC -1- FUNCTION: PROBABLY CONTRIBUTES TO TERMINAL COMPACTATION.
 CC -1- SUBUNIT: HETEROTETRAMER OF TWO TYPE I AND TWO TYPE II KERATINS.
 CC -1- TISSUE SPECIFICITY: IN THE SUPRABASAL LAYERS OF EPIDERMAL
 CC TISSUES FROM MOST BODY SITES, EXCEPT IN FORESKIN, IN SQUAMOUS
 CC METAPLASTS AND CARCINOMAS.
 CC -1- DEVELOPMENTAL STAGE: SYNTHESIZED DURING MATURATION OF EPIDERMAL
 CC KERATINOCYTES.
 CC -1- DISEASE: DEFECTS IN KRT2A ARE A CAUSE OF ICHTHYOSIS BULLOSA OF
 CC SIEMENS (IBS). A RARE AUTOSOMAL DOMINANT DISORDER DISPLAYING A
 CC TYPE OF EPIDERMOLYTIC HYPERKERATOSIS CHARACTERIZED BY EXTENSIVE
 CC BLISTERING FROM BIRTH. HYPERKERATOSIS AND SHEDDING OF THE OUTER
 CC LAYERS OF THE EPIDERMIS (MOLTING) ARE OBSERVED IN LATER WEEKS.
 CC -1- MISCELLANEOUS: THERE ARE TWO TYPES OF CYTOSKELETAL AND
 CC MICROFILAMENTAR KERATIN: I (ACIDIC; 40-55 kDa) [K9 TO K20] AND II
 CC (NEUTRAL TO BASIC; 56-70 kDa) [K1 TO K8].
 CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M99061; AAC83410.1; -;
 DR EMBL; AF019084; AAB81946.1; -;
 DR PIR; A44861; A44861.
 DR Genew; HGNC:6439; KRT2A.
 DR MIM; 600194; -;
 DR MIM; 146800; -;
 DR InterPro; IPR001664; IF.
 DR InterPro; IPR003054; Keratin_II.
 DR Pfam; PF00038; filament; 1.
 DR PRINTS; PR01276; TYPE2KERATIN.
 DR PROSITE; PS00226; IF_1.
 KW Intermediate filament; Coiled coil; Keratin; Disease mutation;
 KW Phosphorylation.
 FT DOMAIN 1
 FT 183
 FT 184 493 HEAD.
 FT 184 494 ROD.
 FT 184 645 TAIL.
 FT 184 219 COIL 1A.
 FT 220 238 LINKER 1.
 FT 239 330 COIL 1B.
 FT 331 354 LINKER 12.
 FT 355 493 COIL 2.
 FT 62 62 PHOSPHORYLATION (BY SIMILARITY).
 FT 187 187 0 -> P (IN IBS).
 FT 187 187 /FTid=VAR_003865.
 FT 188 188 I -> N (IN IBS).
 FT 188 192 /FTid=VAR_010514.
 FT 192 192 N -> D (IN IBS).
 FT 192 192 /FTid=VAR_010515.
 FT 192 192 N -> Y (IN IBS).
 FT 482 482 E -> K (IN IBS).
 FT 482 482 /FTid=VAR_009185.
 FT 485 485 T -> P (IN IBS).
 FT 485 485 /FTid=VAR_009186.
 FT 490 490 L -> P (IN IBS).
 FT 490 490 /FTid=VAR_009187.
 FT 493 493 E -> D (IN IBS).
 FT 493 493 /FTid=VAR_010516.
 FT 493 493 E -> K (IN IBS).
 FT 493 493 /FTid=VAR_003866.
 FT 493 493 E -> K (IN IBS).
 FT 493 493 /FTid=VAR_003867.
 SQ SEQUENCE 645 AA; 65865 MW; EE025A173E33409A CRC64;
 Query Match 100.0%; Score 31; DB 1; Length 645;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 GGFSP 5
 Db 145 GGFSP 149
 RESULT 14
 TR2M_AGRV1 STANDARD; PRT; 723 AA.
 AC 004564;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DE Tryptophan 2-monooxygenase (EC 1.13.12.3).
 GN IAAH.
 OS Agrobacterium vitis (Rhizobium vitis).
 OC Bacteri; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Rhizobium.


```

OX  NCBI_TaxID=373;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=S4;
RX  MEDLINE=93101133; PubMed=1465104;
RA  "Canaday J., Gerard J.-C., Crouzet P., Otten L.;
RT  "Organization and functional analysis of three T-DNAs from the
RT  vtiopline Ti plasmid pRiA4."
RL  Mol. Gen. Genet. 235:292-303(1992).
CC  -1- CATALYTIC ACTIVITY: L-tryptophan + O(2) -> indole-3-acetamide +
CC  CO(2) + H(2)O.
CC  -1- PATHWAY: FIRST STEP IN THE BIOSYNTHESIS OF AUXINS FROM TRYPTOPHAN.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL: M91609; AAA98149.1; -
DR  InterPro: IPR002937; Amino_oxidase.
DR  InterPro: IPR000624; Glycosidase.
DR  InterPro: IPR000205; NAD_binding.
DR  InterPro: IPR001878; ZnF_CCHC.
DR  Pfam: PF01593; Amino_oxidase; 1.
DR  Pfam: PF02027; RoLb_RoLc; 1.
DR  SMART: SM00343; ZnF_C2HC; 1.
KW  Oxidoreductase; Monooxygenase; Auxin biosynthesis; Crown gall tumor;
KW  T-DNA; Plasmid.
SQ  SEQUENCE 723 AA; 80803 MW; ABAC9A94B6B4E36 CRC64;

Query Match          100.0%; Score 31; DB 1; Length 723;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 GGFGP 5
    |||||
DB  410 GGFGP 414

RESULT 15
TR2M_AGRH          STANDARD;      PRT;      749 AA.
AC  Q09109;
DT  01-NOV-1995 (Rel. 32, Created)
DT  01-NOV-1995 (Rel. 32, Last sequence update)
DT  16-OCT-2001 (Rel. 40, Last annotation update)
DE  Tryptophan 2-monooxygenase (EC 1.13.12.3).
GN  AUX1.
OS  Agrobacterium rhizogenes.
OG  Plasmid pRiA4.
OC  Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC  Rhizobiaceae; Rhizobium.
OX  NCBI_TaxID=359;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=A4;
RX  MEDLINE=92033088; PubMed=1932811;
RA  Camilleri C., Jouanin L.;
RT  "The Tr-DNA region carrying the auxin synthesis genes of the
RT  Agrobacterium rhizogenes agroplene-type plasmid pRiA4: nucleotide
RT  sequence analysis and introduction into tobacco plants."
RL  Mol. Plant Microbe Interact. 4:155-162(1991).
CC  -1- CATALYTIC ACTIVITY: L-tryptophan + O(2) -> indole-3-acetamide +
CC  CO(2) + H(2)O.
CC  -1- PATHWAY: FIRST STEP IN THE BIOSYNTHESIS OF AUXINS FROM TRYPTOPHAN.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way

```

```

CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL: M61151; AAA22080.1; -
DR  InterPro: IPR002937; Amino_oxidase.
DR  InterPro: IPR000624; Glycosidase.
DR  InterPro: IPR000205; NAD_binding.
DR  Pfam: PF01593; Amino_oxidase; 1.
DR  Pfam: PF02027; RoLb_RoLc; 1.
KW  Oxidoreductase; Monooxygenase; Auxin biosynthesis; T-DNA; Plasmid.
SQ  SEQUENCE 749 AA; 83178 MW; 8B7793970206F97B CRC64;

Query Match          100.0%; Score 31; DB 1; Length 749;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 GGFGP 5
    |||||
DB  424 GGFGP 428

```

Search completed: March 17, 2003, 12:07:17
 Job time : 10 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 17, 2003, 12:05:48 ; Search time 25 Seconds
(without alignments)
41,209 Million cell updates/sec

Title: SEQ25-FG
Perfect score: 31
Sequence: 1 g9fgp 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPREMBL_21: *
1: sp_archaea: *
2: sp_bacteria: *
3: sp_fungi: *
4: sp_human: *
5: sp_invertebrate: *
6: sp_mammal: *
7: sp_mhc: *
8: sp_organelle: *
9: sp_phage: *
10: sp_plant: *
11: sp_rodent: *
12: sp_virus: *
13: sp_vertebrate: *
14: sp_unclassified: *
15: sp_virus: *
16: sp_bacteriap: *
17: sp_archaeap: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31	100.0	56	17	09YC87
2	31	100.0	62	13	042562
3	31	100.0	74	5	061783
4	31	100.0	76	10	040855
5	31	100.0	93	2	08RIV5
6	31	100.0	121	12	08QPV1
7	31	100.0	124	16	09CLP6
8	31	100.0	130	17	09YFL9
9	31	100.0	132	6	09B662
10	31	100.0	137	10	09A066
11	31	100.0	147	2	09K5H6
12	31	100.0	150	2	09K5H4
13	31	100.0	150	2	09K5H4
14	31	100.0	154	2	09K5H0
15	31	100.0	155	2	09K5H2
16	31	100.0	158	2	09K5G6

17	31	100.0	159	2	09K5G8	09K5G8 bordetella
18	31	100.0	159	2	09ALP4	09ALP4 bordetella
19	31	100.0	160	2	09K5I0	09K5I0 bordetella
20	31	100.0	160	2	09K5G3	09K5G3 bordetella
21	31	100.0	161	2	09K5G2	09K5G2 bordetella
22	31	100.0	167	2	09J7U5	09J7U5 bordetella
23	31	100.0	169	16	08UGM7	08UGM7 bordetella
24	31	100.0	173	2	09ALQ5	09ALQ5 bordetella
25	31	100.0	177	2	09ALQ3	09ALQ3 bordetella
26	31	100.0	178	2	09J3T1	09J3T1 bordetella
27	31	100.0	178	10	09M8R5	09M8R5 arabidopsis
28	31	100.0	179	2	09AHP1	09AHP1 bordetella
29	31	100.0	181	2	09ALQ1	09ALQ1 bordetella
30	31	100.0	184	2	09ALP6	09ALP6 bordetella
31	31	100.0	192	16	09RKN2	09RKN2 streptomyces
32	31	100.0	195	2	09ALP8	09ALP8 bordetella
33	31	100.0	197	16	09F2V0	09F2V0 streptomyces
34	31	100.0	206	2	09KJY2	09KJY2 bordetella
35	31	100.0	211	2	09KJY3	09KJY3 bordetella
36	31	100.0	212	2	09APU1	09APU1 bordetella
37	31	100.0	212	5	09VBD1	09VBD1 drosophila
38	31	100.0	216	2	09KJY4	09KJY4 bordetella
39	31	100.0	225	10	064772	064772 arabidopsis
40	31	100.0	233	4	09GCV2	09GCV2 homo sapien
41	31	100.0	235	4	08MWC3	08MWC3 homo sapien
42	31	100.0	235	5	002442	002442 globodera r
43	31	100.0	238	16	09CIV5	09CIV5 lactococcus
44	31	100.0	250	13	09PUY2	09PUY2 apteronotus
45	31	100.0	266	4	08WY24	08WY24 homo sapien

ALIGNMENTS

RESULT 1
ID 09YC87 PRELIMINARY: PRT: 56 AA.
AC 09YC87;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DT Hypothetical protein APE5051.
GN APE5051.
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcaceae;
OC Desulfurococcaceae; Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K1;
RX MEDLINE=99310339; PubMed=10382966;
RA KAWARABAYASHI Y., HINO Y., HORIKAWA H., YAMAZAKI S., HAIKAWA Y.,
RA Jin-no K., TAKAHASHI M., SEKINE M., BABA S.-I., ANKAI A., KOSUGI H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kishida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.,
RT "Complete genome sequence of an aerobic hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix K1.";
RL DNA Res. 6:83-101(1999).
DR EMBL: AP000061; BAA80361.1;
KW Hypothetical protein; Complete proteome.
SO SEQUENCE 56 AA; 6118 MW; 35DAD5A598ECP3A2 CRC64;

Query Match 100.0%; Score 31; DB 17; Length 56;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGRFP 5
| | | | |
DB 42 GGRFP 46

RESULT 2
ID 042562 PRELIMINARY; PRT; 62 AA.
AC 042562;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE Hydroxybutyrate dehydrogenase.
GN FRHBD.
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Takifugu.
OX NCBI_TaxID=31033;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97285126; PubMed=9140399;
RA Aparicio S., Hawker K., Collage A., Mikawa Y., Zuo L., Venkatesh B.,
RT Chen E., Krumlauf R., Brenner S.;
RT "Organization of the Fugu rubripes Hox clusters: evidence for
RL continuing evolution of vertebrate Hox complexes.";
RL Nat. Genet. 16:79-83(1997).
DR EMBL: U95612; AAB69987.1; -
SQ SEQUENCE 62 AA; 6498 MW; 82189D12C95F3CDD CRC64;

Query Match 100.0%; Score 31; DB 13; Length 62;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGF 5
Db 30 GGF 34

RESULT 3
ID 061783 PRELIMINARY; PRT; 74 AA.
AC 061783;
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Hypothetical 7.4 kDa protein.
GN R12E2.6.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodermidae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RL investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Goella D., Scheet P.;
RT "The sequence of C. elegans cosmid R12E2.";
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF067219; AAC17027.1; -
KW Hypothetical protein.
SQ SEQUENCE 74 AA; 7380 MW; D4097BFF9F3A176E CRC64;

Query Match 100.0%; Score 31; DB 5; Length 74;

Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGF 5
Db 45 GGF 49

RESULT 4
ID 040855 PRELIMINARY; PRT; 76 AA.
AC 040855;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Glycine-rich cell wall protein.
GN EMB31.
OS Picea glauca (White spruce).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Picea.
OX NCBI_TaxID=3330;
RN [1]
RP SEQUENCE FROM N.A.
RA Dong J.-Z., Dunstan D.I.;
RT "Gene expression during somatic embryogenesis.";
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: L47748; AAB01565.1; -
SQ SEQUENCE 76 AA; 7242 MW; 9B5B3E46DA6E3F70 CRC64;

Query Match 100.0%; Score 31; DB 10; Length 76;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGF 5
Db 26 GGF 30

RESULT 5
ID 08R1V5 PRELIMINARY; PRT; 93 AA.
AC 08R1V5;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Paraben-hydrolyzing esterase (Fragment).
GN PRBA.
OS Enterobacter gergoviae.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Enterobacter.
OX NCBI_TaxID=61647;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-G1, AND G12;
RA Valkova N., Lepine F., Bollet C., Dupont M., Villemur R.;
RT "Prba, a gene coding for an esterase hydrolyzing parabens in E.
RL cloacae and E. gergoviae strains.";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY077722; AAL82803.1; -
DR EMBL: AY077723; AAL82804.1; -
FT NON_TER
FT NON_TER
SQ SEQUENCE 93 AA; 10149 MW; F61A94CE6BA46556 CRC64;

Query Match 100.0%; Score 31; DB 2; Length 93;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGF 5
Db 56 GGF 60

RESULT 6

OBOPV1 PRELIMINARY; PRT; 121 AA.
 AC OBOPV1; 01-JUN-2002 (TREMblrel. 21, Created)
 DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)
 DE 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Capsid protein (Fragment).
 GN AV1.
 OS tomato chlorotic vein virus.
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
 OX NCBI_TaxID=172390;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DF-BRL.
 RA Ribeiro S.G., Ambrozewicz L.P., de Avila A.C., Galegario R.F.,
 RA Fernandes J.J., Lima M.F., Meilo R.N., Rocha H., Zerbini F.M.,
 RT "Distribution and genetic diversity of tomato-infecting geminiviruses
 RT in Brazil.";
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY049205; AAL82823.1;
 FT NON_TER 121 121
 SQ SEQUENCE 121 AA; 13806 MW; 89D051C65FAE2AF6 CRC64;

Query Match 100.0%; Score 31; DB 12; Length 121;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 GGFGP 5
 Db 27 GGFGP 31

RESULT 7

O9CLP6 PRELIMINARY; PRT; 124 AA.
 AC O9CLP6; 01-JUN-2001 (TREMblrel. 17, Created)
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
 DE 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE Hypothetical protein PM1171.
 GN PM1171.
 OS Pasteurella multocida.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Pasteurella.
 OX NCBI_TaxID=747;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PM70.
 RA MEDLINE=21145866; PubMed=11248100;
 RA May B.J., Zhang Q., Li L.L., Pauslan M.L., Whittam T.S., Kapur V.,
 RT "Complete genomic sequence of Pasteurella multocida pm0.",
 RT Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
 DR EMBL: AEO06157; MAK03255.1;
 KM Hypothetical protein; Complete proteome.
 SQ SEQUENCE 124 AA; 13402 MW; 289E9736A09BBAA4 CRC64;

Query Match 100.0%; Score 31; DB 16; Length 124;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 GGFGP 5
 Db 33 GGFGP 37

RESULT 8

O9YFL9 PRELIMINARY; PRT; 130 AA.
 AC O9YFL9; 01-NOV-1999 (TREMblrel. 12, Created)
 DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)
 DE 01-MAR-2002 (TREMblrel. 20, Last annotation update)

DE Hypothetical protein APE0229.
 GN APE0229.
 OS Aeropyrum pernix.
 OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcates;
 OC Desulfurococcaceae; Aeropyrium.
 OX NCBI_TaxID=56636;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KI;
 RX MEDLINE=99310339; PubMed=10382966;
 RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Halkawa Y.,
 RA Jin-no K., Takehashi M., Sekine M., Baba S.-I., Anai A., Kosugi H.,
 RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
 RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kuoh Y.,
 RA Yamazaki J., Kushiida N., Oguchi A., Aoki K.-I., Kubota K.,
 RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.,
 RT "Complete genome sequence of an aerobic hyper-thermophilic
 RT crenarchaeon, Aeropyrum pernix K1.";
 RL DNA Res. 6:83-101(1999).
 DR EMBL: AP000058; BAA79142.1;
 KM Hypothetical protein; Complete proteome.
 SQ SEQUENCE 130 AA; 13829 MW; F388BA63EB57729F CRC64;

Query Match 100.0%; Score 31; DB 17; Length 130;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 GGFGP 5
 Db 61 GGFGP 65

RESULT 9

O9BG62 PRELIMINARY; PRT; 132 AA.
 AC O9BG62; 01-JUN-2001 (TREMblrel. 17, Created)
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
 DE 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE Heparin-binding epidermal growth factor (Fragment).
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Klonisch T., Wolf P., Hombach-Klonisch S., Vogt S., Kuechenhoff A.,
 RA Tetens F., Fischer B.,
 RT "Erdb genes and epidermal growth factor- (EGF-) like ligands in the
 RT perit-implantation rabbit uterus and blastocyst.";
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF333182; MAK14375.1;
 DR HSSP: Q99075; 1XDT.
 DR Interpro: IPR000561; EGF_1like.
 DR Pfam: PF00008; EGF_1.
 DR PRINTS: PR00009; EGFTEF.
 DR SMART: SM00181; EGF_1.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
 DR PROSITE: PS01186; EGF_2; 1.
 KW EGF-like domain; Glycoprotein.
 FT NON_TER 1 132
 SQ SEQUENCE 132 AA; 14628 MW; 6F5ACC07E6A0FF4E CRC64;

Query Match 100.0%; Score 31; DB 6; Length 132;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 GGFGP 5
 Db 11 GGFGP 15

RESULT 10

Q9AUG6 PRELIMINARY; PRT; 137 AA.
 AC Q9AUG6; 09AUG66;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Malate synthase (EC 4.1.3.2) (Fragment).
 OS Typha latifolia (Common cattail).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Typhaceae; Typha.
 NCBI_TaxID=4733;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21293323; PubMed=11399149;
 RA Lewis C.E., Doyle J.J.;
 RT "Phylogenetic utility of the nuclear gene malate synthase in the palm
 family (Arecaceae)."
 RL Mol. Phylogenet. Evol. 19:409-420(2001).
 CC -1- CATALYTIC ACTIVITY: L-MALATE + COA = ACETYL-COA + H(2)O +
 CC GLYOXYLATE.
 CC -1- PATHWAY: SECOND STEP IN GLYOXYLATE BYPASS, AN ALTERNATIVE TO THE
 CC TRICARBOXYLIC ACID CYCLE (IN BACTERIA, FUNGI AND PLANTS).
 CC -1- SIMILARITY: BELONGS TO THE MALATE SYNTHASE FAMILY.
 DR EMBL: AF249966; AK28015.1; -
 DR InterPro: IPR001465; Malate_synthase.
 DR Pfam: PF01274; Malate_synthase; 1.
 DR PROSITE: PS00510; MALATE_SYNTHASE; 1.
 DR Glyoxylate bypass; Lyase; Tricarboxylic acid cycle.
 FT NON_TER 1 1
 FT SEQUENCE 137 AA; 15721 MW; F6093DB80C61B8A9 CRC64;

Query Match 100.0%; Score 31; DB 10; Length 137;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFGP 5
 DB 56 GGFGP 60

RESULT 11

Q9K5H6 PRELIMINARY; PRT; 147 AA.
 AC Q9K5H6; 09K5H6;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE Pertactin (P.68) (Fragment).
 OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
 OC Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;
 OC Bordetella.
 NCBI_TaxID=518;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=5;
 RX MEDLINE=20359389; PubMed=10899896;
 RA Boursaux-Eude C., Guiso N.;
 RT "Polymorphism of Repeated Regions of Pertactin in Bordetella
 pertussis, Bordetella parapertussis, and Bordetella bronchiseptica."
 RL Infect. Immun. 68:4815-4817(2000).
 DR EMBL: AJ250080; CAB76434.1; -
 FT NON_TER 1 1
 FT SEQUENCE 147 AA; 14122 MW; 83F8E6526202E5E3 CRC64;

Query Match 100.0%; Score 31; DB 2; Length 147;
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFGP 5
 DB 69 GGFGP 73

RESULT 12

Q9K5H8 PRELIMINARY; PRT; 150 AA.
 AC Q9K5H8; 09K5H8;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE Pertactin (P.68) (Fragment).
 OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
 OC Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;
 OC Bordetella.
 NCBI_TaxID=518;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LAPR;
 RX MEDLINE=20359389; PubMed=10899896;
 RA Boursaux-Eude C., Guiso N.;
 RT "Polymorphism of Repeated Regions of Pertactin in Bordetella
 pertussis, Bordetella parapertussis, and Bordetella bronchiseptica."
 RL Infect. Immun. 68:4815-4817(2000).
 DR EMBL: AJ250078; CAB76432.1; -
 FT NON_TER 1 1
 FT SEQUENCE 150 AA; 14614 MW; B1CECC7FD4A16725 CRC64;

Query Match 100.0%; Score 31; DB 2; Length 150;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFGP 5
 DB 75 GGFGP 79

RESULT 13

Q9K5H4 PRELIMINARY; PRT; 150 AA.
 AC Q9K5H4; 09K5H4;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE Pertactin (P.68) (Fragment).
 OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
 OC Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;
 OC Bordetella.
 NCBI_TaxID=518;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=335;
 RX MEDLINE=20359389; PubMed=10899896;
 RA Boursaux-Eude C., Guiso N.;
 RT "Polymorphism of Repeated Regions of Pertactin in Bordetella
 pertussis, Bordetella parapertussis, and Bordetella bronchiseptica."
 RL Infect. Immun. 68:4815-4817(2000).
 DR EMBL: AJ250082; CAB76436.1; -
 FT NON_TER 1 1
 FT SEQUENCE 150 AA; 14425 MW; C3AA820B3CEAEAC0 CRC64;

Query Match 100.0%; Score 31; DB 2; Length 150;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFGP 5
 DB 72 GGFGP 76

Job time : 28 secs

RESULT 14

ID Q9K5H0 PRELIMINARY; PRT; 154 AA.
 AC Q9K5H0;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE Pertactin (P.68) (Fragment).
 GN PRN.
 OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
 OC Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;
 OC Bordetella.
 OX NCBI_TaxID=518;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BBCH;
 RX MEDLINE=20359389; PubMed=10899896;
 RA Boursaux-Eude C., Guiso N.;
 RT "Polymorphism of Repeated Regions of Pertactin in Bordetella
 pertussis, Bordetella parapertussis, and Bordetella bronchiseptica."
 RL Infect. Immun. 68:4815-4817(2000).
 DR EMBL; AJ250086; CAB76440.1;
 FT NON_TER 1
 FT NON_TER 1
 FT NON_TER 154
 SQ SEQUENCE 154 AA; 14799 MW; 9D28B10C4249D23E CRC64;

Query Match

Best Local Similarity 100.0%; Score 31; DB 2; Length 154;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFGP 5
 |||||
 Db 75 GGFGP 79

RESULT 15

ID Q9K5H2 PRELIMINARY; PRT; 155 AA.
 AC Q9K5H2;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE Pertactin (P.68) (Fragment).
 GN PRN.
 OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
 OC Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;
 OC Bordetella.
 OX NCBI_TaxID=518;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CVGEQ;
 RX MEDLINE=20359389; PubMed=10899896;
 RA Boursaux-Eude C., Guiso N.;
 RT "Polymorphism of Repeated Regions of Pertactin in Bordetella
 pertussis, Bordetella parapertussis, and Bordetella bronchiseptica."
 RL Infect. Immun. 68:4815-4817(2000).
 DR EMBL; AJ250084; CAB76438.1;
 FT NON_TER 1
 FT NON_TER 1
 FT NON_TER 155
 SQ SEQUENCE 155 AA; 14899 MW; E8A5CF447F8CE43E CRC64;

Query Match

Best Local Similarity 100.0%; Score 31; DB 2; Length 155;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFGP 5
 |||||
 Db 76 GGFGP 80

Search completed: March 17, 2003, 12:10:21

GenCore version 5.1.4-P5-4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 17, 2003, 12:07:24 ; Search time 152 Seconds
(without alignments)
21.208 Million cell updates/sec

Title: SEQ25-FG
Perfect score: 31
Sequence: 1 gfgfp 5

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_AA_Main:*

1: /cgn2_6/ptodata/2/paa/US06.COMB.pep:*
2: /cgn2_6/ptodata/2/paa/US06.COMB.pep:*
3: /cgn2_6/ptodata/2/paa/US07.COMB.pep:*
4: /cgn2_6/ptodata/2/paa/US08.COMB.pep:*
5: /cgn2_6/ptodata/2/paa/US08.COMB.pep:*
6: /cgn2_6/ptodata/2/paa/US08.COMB.pep:*
7: /cgn2_6/ptodata/2/paa/US08.COMB.pep:*
8: /cgn2_6/ptodata/2/paa/US08.COMB.pep:*
9: /cgn2_6/ptodata/2/paa/US08.COMB.pep:*
10: /cgn2_6/ptodata/2/paa/US08.COMB.pep:*
11: /cgn2_6/ptodata/2/paa/US08.COMB.pep:*
12: /cgn2_6/ptodata/2/paa/US08.COMB.pep:*
13: /cgn2_6/ptodata/2/paa/US08.COMB.pep:*
14: /cgn2_6/ptodata/2/paa/US09.COMB.pep:*
15: /cgn2_6/ptodata/2/paa/US09.COMB.pep:*
16: /cgn2_6/ptodata/2/paa/US09.COMB.pep:*
17: /cgn2_6/ptodata/2/paa/US09.COMB.pep:*
18: /cgn2_6/ptodata/2/paa/US09.COMB.pep:*
19: /cgn2_6/ptodata/2/paa/US09.COMB.pep:*
20: /cgn2_6/ptodata/2/paa/US09.COMB.pep:*
21: /cgn2_6/ptodata/2/paa/US09.COMB.pep:*
22: /cgn2_6/ptodata/2/paa/US09.COMB.pep:*
23: /cgn2_6/ptodata/2/paa/US09.COMB.pep:*
24: /cgn2_6/ptodata/2/paa/US10.COMB.pep:*
25: /cgn2_6/ptodata/2/paa/US10.COMB.pep:*
26: /cgn2_6/ptodata/2/paa/US10.COMB.pep:*
27: /cgn2_6/ptodata/2/paa/US60.COMB.pep:*

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	31	100.0	26	1	PCT-US01-05614-625
2	31	100.0	26	16	US-09-205-258-625
3	31	100.0	26	23	US-09-933-767-625
4	31	100.0	26	24	US-10-004-860-625
5	31	100.0	26	24	US-10-023-282-625
6	31	100.0	39	27	US-60-163-233-2462

7	31	100.0	46	22	US-09-855-754-8	Sequence 8, Appl
8	31	100.0	47	17	US-09-307-590-438	Sequence 438, App
9	31	100.0	47	21	US-09-708-427-74924	Sequence 74924, A
10	31	100.0	47	22	US-09-855-548-438	Sequence 438, App
11	31	100.0	51	22	US-09-855-754-7	Sequence 7, Appl
12	31	100.0	51	22	US-09-855-754-13	Sequence 13, Appl
13	31	100.0	56	22	US-09-855-754-9	Sequence 9, Appl
14	31	100.0	56	22	US-09-855-754-10	Sequence 12, Appl
15	31	100.0	56	22	US-09-855-754-12	Sequence 3922, Ap
16	31	100.0	60	27	US-60-192-739-3922	Sequence 2835, Ap
17	31	100.0	61	22	US-09-855-754-11	Sequence 11, Appl
18	31	100.0	63	17	US-09-330-781-278	Sequence 278, App
19	31	100.0	63	22	US-09-828-769-278	Sequence 278, App
20	31	100.0	66	27	US-60-160-189-7443	Sequence 7443, App
21	31	100.0	66	27	US-60-195-311-575	Sequence 575, App
22	31	100.0	69	27	US-60-169-867-7578	Sequence 7578, App
23	31	100.0	78	24	US-10-015-127-10029	Sequence 10029, A
24	31	100.0	84	27	US-60-196-718-6717	Sequence 6717, Ap
25	31	100.0	86	1	PCT-US99-228538-57	Sequence 57, Appl
26	31	100.0	87	20	US-09-617-682A-3729	Sequence 3729, Ap
27	31	100.0	90	20	US-09-617-682A-3728	Sequence 3728, Ap
28	31	100.0	96	26	US-10-221-279-10878	Sequence 10878, A
29	31	100.0	97	1	PCT-US99-228538-56	Sequence 56, Appl
30	31	100.0	97	26	US-10-221-279-9310	Sequence 9310, Ap
31	31	100.0	99	24	US-10-013-127-12029	Sequence 12029, A
32	31	100.0	99	26	US-10-221-279-7115	Sequence 7115, Ap
33	31	100.0	108	26	US-09-649-866A-2533	Sequence 2533, Ap
34	31	100.0	109	20	US-09-649-866A-2537	Sequence 2537, Ap
35	31	100.0	113	20	US-09-620-111B-2337	Sequence 2337, Ap
36	31	100.0	113	26	US-10-221-279-6341	Sequence 6341, Ap
37	31	100.0	113	27	US-60-162-247-2815	Sequence 2815, Ap
38	31	100.0	113	27	US-60-169-840-6988	Sequence 6988, Ap
39	31	100.0	113	27	US-60-169-867-6126	Sequence 6126, Ap
40	31	100.0	114	21	US-09-708-427-56523	Sequence 56523, A
41	31	100.0	114	21	US-09-708-427-76747	Sequence 76747, A
42	31	100.0	114	21	US-09-708-427-76747	Sequence 44, Appl
43	31	100.0	119	21	US-09-760-461-44	Sequence 659, Appl
44	31	100.0	119	21	US-09-760-461-44	Sequence 44, Appl
45	31	100.0	119	25	US-10-144-132-44	

ALIGNMENTS

RESULT 1
PCT-US01-05614-625
Sequence 625, Application PC/TUS0105614
GENERAL INFORMATION:
APPLICANT: Human Genome Sciences, Inc.
TITLE OF INVENTION: 207 Human Secreted Proteins
FILE REFERENCE: P2007PCW2
CURRENT APPLICATION NUMBER: PCT/US01/05614
CURRENT FILING DATE: 2001-02-21
PRIOR APPLICATION NUMBER: 60/184,836
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: 60/193,170
PRIOR FILING DATE: 2000-03-29
NUMBER OF SEQ ID NOS: 1245
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 625
LENGTH: 26
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US01-05614-625

Query Match 100.0%; Score 31; DB 1; Length 26;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GFGFP 5
DB 18 GFGFP 22

```
RESULT 2
US-09-205-258-625
; Sequence 625, Application US/09205258
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: P2007P1
; CURRENT APPLICATION NUMBER: US/09/205,258
; CURRENT FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
```

```
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 625
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-205-258-625
```

```
Query Match      100.0%; Score 31; DB 16; Length 26;
Best Local Similarity 100.0%; Pred. No. 2,6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 GGRGP 5
Db      18 GGRGP 22
```

```
RESULT 3
US-09-933-767-625
; Sequence 625, Application US/09933767
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: P2007P2
; CURRENT APPLICATION NUMBER: US/09/933,767
; CURRENT FILING DATE: 2001-08-22
; EARLIER APPLICATION NUMBER: PCT/US01/05614
; EARLIER FILING DATE: 2001-02-21
; EARLIER APPLICATION NUMBER: 60/184,836
; EARLIER FILING DATE: 2000-02-24
; EARLIER APPLICATION NUMBER: 60/193,170
; EARLIER FILING DATE: 2000-03-29
; EARLIER APPLICATION NUMBER: 09/205,258
; EARLIER FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
```

PRIOR APPLICATION NUMBER: 60/048,971
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,964
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,882
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,899
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,893
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,900
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,901
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,892
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,915
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/049,019
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,970
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,972
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/049,373
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,875
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/049,374
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,917
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,949
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,974
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,883
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,897
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,898
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,962
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,963
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,877
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,878
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/068,054
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/068,064
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/068,053
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/070,923
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/073,160
PRIOR FILING DATE: 1998-01-30
PRIOR APPLICATION NUMBER: 60/073,159
PRIOR FILING DATE: 1998-01-30
PRIOR APPLICATION NUMBER: 60/073,165
PRIOR FILING DATE: 1998-01-30
PRIOR APPLICATION NUMBER: 60/073,164
PRIOR FILING DATE: 1998-01-30
PRIOR APPLICATION NUMBER: 60/085,925
PRIOR FILING DATE: 1998-05-18
PRIOR APPLICATION NUMBER: 60/085,921
PRIOR FILING DATE: 1998-05-18
PRIOR APPLICATION NUMBER: 60/085,923

PRIOR FILING DATE: 1998-05-18
PRIOR APPLICATION NUMBER: 60/085,922
PRIOR FILING DATE: 1998-05-18
PRIOR APPLICATION NUMBER: 60/092,921
PRIOR FILING DATE: 1998-07-15
PRIOR APPLICATION NUMBER: 60/094,657
PRIOR FILING DATE: 1998-07-30
NUMBER OF SEQ ID NOS: 1245
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 625
LENGTH: 26
TYPE: PRT
ORGANISM: Homo sapiens
US-09-933-767-625

Query Match
Best Local Similarity 100.0%; Score 31; DB 23; Length 26;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFEP 5
DB 18 GGFEP 22

RESULT 4
US-10-004-860-625
Sequence 625, Application US/10004860
GENERAL INFORMATION:
APPLICANT: Young et al.
TITLE OF INVENTION: 207 Human Secreted Proteins
FILE REFERENCE: P2007P1
CURRENT APPLICATION NUMBER: US/10/004,860
CURRENT FILING DATE: 2001-12-07
Prior Application removed - See file wrapper or Palm
NUMBER OF SEQ ID NOS: 1227
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 625
LENGTH: 26
TYPE: PRT
ORGANISM: Homo sapiens
US-10-004-860-625

Query Match
Best Local Similarity 100.0%; Score 31; DB 24; Length 26;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFEP 5
DB 18 GGFEP 22

RESULT 5
US-10-023-282-625
Sequence 625, Application US/10023282
GENERAL INFORMATION:
APPLICANT: Young et al.
TITLE OF INVENTION: 207 Human Secreted Proteins
FILE REFERENCE: P2007P1
CURRENT APPLICATION NUMBER: US/10/023,282
CURRENT FILING DATE: 2001-12-20
EARLIER APPLICATION NUMBER: 09/205,258
EARLIER FILING DATE: 1998-12-04
EARLIER APPLICATION NUMBER: PCT/US98/11422
EARLIER FILING DATE: 1998-06-04
EARLIER APPLICATION NUMBER: 60/048,885
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,375
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,881
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,880
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,896

```

; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 625
; LENGTH: 26

```

```

; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-023-282-625
;
; Query Match
; Best Local Similarity 100.0%; Score 31; DB 24; Length 26;
; Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 1 GGFPG 5
; DB 18 GGFPG 22
;
; RESULT 6
; US-60-163-233-2462
; Sequence 2462, Application US/60163233
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN DRUG TARGET PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN DRUG TARGET PROTEINS
; FILE REFERENCE: C100013
; CURRENT APPLICATION NUMBER: US/60/163,233
; CURRENT FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 4426
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2462
; LENGTH: 39
; TYPE: PRT
; ORGANISM: Human
; US-60-163-233-2462
;
; Query Match
; Best Local Similarity 100.0%; Score 31; DB 27; Length 39;
; Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 1 GGFPG 5
; DB 34 GGFPG 38
;
; RESULT 7
; US-09-855-754-8
; Sequence 8, Application US/09855754
; GENERAL INFORMATION:
; APPLICANT: BOURSAUX-EUDE, CAROLINE
; TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED
; TITLE OF INVENTION: REGIONS OF PERTACTIN IN BORDETELLA PERTUSSIS,
; TITLE OF INVENTION: BORDETELLA PARAPERTUSSIS, AND BORDETELLA
; TITLE OF INVENTION: BRONCHISEPTICA, THEIR USE IN DIAGNOSTICS, AND IN
; FILE REFERENCE: 03495-0206-00000
; CURRENT APPLICATION NUMBER: US/09/855,754
; CURRENT FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/206,969
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 8
; LENGTH: 46
; TYPE: PRT
; ORGANISM: Bordetella bronchiseptica
; US-09-855-754-8
;
; Query Match
; Best Local Similarity 100.0%; Score 31; DB 22; Length 46;
; Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 1 GGFPG 5
; DB 23 GGFPG 27

```

RESULT 8

US-09-307-590-438
Sequence 438, Application US/09307590
GENERAL INFORMATION:
APPLICANT: Culpepper, Janice A.
APPLICANT: Gearling, David P.
APPLICANT: Holtzman, Douglas A.
APPLICANT: Robison, Keith E.
TITLE OF INVENTION: NOCLETIC ACID MOLECULES DERIVED FROM A
FILE REFERENCE: MLN98-12PA
CURRENT APPLICATION NUMBER: US/09/307,590
PRIOR FILING DATE: 1999-05-07
EARLIER APPLICATION NUMBER: 60/084,683
PRIOR FILING DATE: 1998-05-08
NUMBER OF SEQ ID NOS: 756
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 438
LENGTH: 47
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SIGNAL
LOCATION: (1)...(20)
NAME/KEY: VARIANT
LOCATION: (1)...(47)
OTHER INFORMATION: Xaa - Any Amino Acid
US-09-307-590-438

Query Match

Best Local Similarity 100.0%; Score 31; DB 17; Length 47;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFGP 5
|||||

DB 42 GGFGP 46

RESULT 9

US-09-708-427-74924
Sequence 74924, Application US/09708427
GENERAL INFORMATION:
APPLICANT: N. ALEXANDROV et al.
TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
FILE REFERENCE: 2750-1243P
CURRENT APPLICATION NUMBER: US/09/708,427
PRIOR FILING DATE: 2000-11-09
NUMBER OF SEQ ID NOS: 85364
SOFTWARE: PatentIn version 3.1
SEQ ID NO 74924
LENGTH: 47
TYPE: PRT
ORGANISM: zea mays subsp. mays
FEATURE:
NAME/KEY: misc-feature
LOCATION: 1..47
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc-feature
LOCATION: 1..47
OTHER INFORMATION: Cereb Seq. ID 1946054
US-09-708-427-74924

Query Match

Best Local Similarity 100.0%; Score 31; DB 21; Length 47;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFGP 5
|||||

DB 3 GGFGP 7

RESULT 10

US-09-855-548-438
Sequence 438, Application US/09855548
GENERAL INFORMATION:
APPLICANT: Culpepper, Janice A.
APPLICANT: Gearling, David P.
APPLICANT: Holtzman, Douglas A.
APPLICANT: Robison, Keith E.
TITLE OF INVENTION: NOCLETIC ACID MOLECULES DERIVED FROM A
FILE REFERENCE: MLN98-12PA
CURRENT APPLICATION NUMBER: US/09/855,548
PRIOR FILING DATE: 2001-05-15
PRIOR APPLICATION NUMBER: 09/307,590
PRIOR FILING DATE: 1999-05-07
PRIOR APPLICATION NUMBER: 60/084,683
PRIOR FILING DATE: 1998-05-08
NUMBER OF SEQ ID NOS: 756
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 438
LENGTH: 47
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SIGNAL
LOCATION: (1)...(20)
NAME/KEY: VARIANT
LOCATION: (1)...(47)
OTHER INFORMATION: Xaa - Any Amino Acid
US-09-855-548-438

Query Match

Best Local Similarity 100.0%; Score 31; DB 22; Length 47;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFGP 5
|||||

DB 42 GGFGP 46

RESULT 11

US-09-855-754-7
Sequence 7, Application US/09855754
GENERAL INFORMATION:
APPLICANT: BOIRSAX-ENDE, CAROLINE
TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED
TITLE OF INVENTION: REGIONS OF PERTACTIN IN BORDETELLA PERTUSSIS,
TITLE OF INVENTION: BORDETELLA PARAPERTUSSIS, AND BORDETELLA
TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS
FILE REFERENCE: 03495-0206-00000
CURRENT APPLICATION NUMBER: US/09/855,754
PRIOR FILING DATE: 2001-09-10
PRIOR APPLICATION NUMBER: 60/206,969
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 7
LENGTH: 51
TYPE: PRT
ORGANISM: Bordetella bronchiseptica
US-09-855-754-7

Query Match

Best Local Similarity 100.0%; Score 31; DB 22; Length 51;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFGP 5
|||||

DB 28 GGFGP 32

RESULT 12
US-09-855-754-13
; Sequence 13, Application US/09855754
; GENERAL INFORMATION:
; APPLICANT: BOURSAX-EUDE, CAROLINE
; APPLICANT: GUISSO-MACLOUF, NICOLE
; TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED
; TITLE OF INVENTION: REGIONS OF PERTACTIN IN BORDETELLA PERTUSSIS,
; TITLE OF INVENTION: BORDETELLA PARAPERTUSSIS, AND BORDETELLA
; TITLE OF INVENTION: BRONCHISEPTICA, THEIR USE IN DIAGNOSTICS, AND IN
; FILE REFERENCE: 03495-0206-00000
; CURRENT APPLICATION NUMBER: US/09/855,754
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/206,969
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 51
; TYPE: PRT
; ORGANISM: Bordetella bronchiseptica
US-09-855-754-13

Query Match
Best Local Similarity 100.0%; Score 31; DB 22; Length 51;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGFEP 5
DB 23 GGFEP 27

RESULT 13
US-09-855-754-9
; Sequence 9, Application US/09855754
; GENERAL INFORMATION:
; APPLICANT: BOURSAX-EUDE, CAROLINE
; APPLICANT: GUISSO-MACLOUF, NICOLE
; TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED
; TITLE OF INVENTION: REGIONS OF PERTACTIN IN BORDETELLA PERTUSSIS,
; TITLE OF INVENTION: BORDETELLA PARAPERTUSSIS, AND BORDETELLA
; TITLE OF INVENTION: BRONCHISEPTICA, THEIR USE IN DIAGNOSTICS, AND IN
; FILE REFERENCE: 03495-0206-00000
; CURRENT APPLICATION NUMBER: US/09/855,754
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/206,969
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Bordetella bronchiseptica
US-09-855-754-9

Query Match
Best Local Similarity 100.0%; Score 31; DB 22; Length 56;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGFEP 5
DB 28 GGFEP 32

RESULT 14
US-09-855-754-10
; Sequence 10, Application US/09855754
; GENERAL INFORMATION:
; APPLICANT: BOURSAX-EUDE, CAROLINE
; APPLICANT: GUISSO-MACLOUF, NICOLE
; TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED

;; TITLE OF INVENTION: REGIONS OF PERTACTIN IN BORDETELLA PERTUSSIS,
;; TITLE OF INVENTION: BORDETELLA PARAPERTUSSIS, AND BORDETELLA
;; TITLE OF INVENTION: BRONCHISEPTICA, THEIR USE IN DIAGNOSTICS, AND IN
;; FILE REFERENCE: 03495-0206-00000
;; CURRENT APPLICATION NUMBER: US/09/855,754
;; PRIOR FILING DATE: 2001-09-10
;; PRIOR APPLICATION NUMBER: 60/206,969
;; PRIOR FILING DATE: 2000-05-25
;; NUMBER OF SEQ ID NOS: 24
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 10
;; LENGTH: 56
;; TYPE: PRT
;; ORGANISM: Bordetella bronchiseptica
US-09-855-754-10

Query Match
Best Local Similarity 100.0%; Score 31; DB 22; Length 56;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFEP 5
DB 28 GGFEP 32

RESULT 15
US-09-855-754-12
; Sequence 12, Application US/09855754
; GENERAL INFORMATION:
; APPLICANT: BOURSAX-EUDE, CAROLINE
; APPLICANT: GUISSO-MACLOUF, NICOLE
; TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED
; TITLE OF INVENTION: REGIONS OF PERTACTIN IN BORDETELLA PERTUSSIS,
; TITLE OF INVENTION: BORDETELLA PARAPERTUSSIS, AND BORDETELLA
; TITLE OF INVENTION: BRONCHISEPTICA, THEIR USE IN DIAGNOSTICS, AND IN
; FILE REFERENCE: 03495-0206-00000
; CURRENT APPLICATION NUMBER: US/09/855,754
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/206,969
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Bordetella bronchiseptica
US-09-855-754-12

Query Match
Best Local Similarity 100.0%; Score 31; DB 22; Length 56;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFEP 5
DB 23 GGFEP 27

Search completed: March 17, 2003, 12:19:36
Job time : 153 secs

NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 7
LENGTH: 51
TYPE: PRT
ORGANISM: Bordetella bronchiseptica
US-09-855-754B-7

Query Match 100.0%; Score 31; DB 5; Length 51;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFGP 5
|||||
DB 28 GGFGP 32

RESULT 3
US-09-855-754B-13
Sequence 13, Application US/09855754B
GENERAL INFORMATION:
APPLICANT: BOURSAX-EDDE, CAROLINE
TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED
TITLE OF INVENTION: REGIONS OF PERTACTIN IN BORDETELLA PERTUSSIS,
TITLE OF INVENTION: BORDETELLA PARAPERTUSSIS, AND BORDETELLA
TITLE OF INVENTION: BRONCHISEPTICA, THEIR USE IN DIAGNOSTICS, AND IN
FILE REFERENCE: 03495-0206-00000
CURRENT APPLICATION NUMBER: US/09/855,754B
CURRENT FILING DATE: 2001-05-16
PRIOR APPLICATION NUMBER: 60/206,969
PRIOR FILING DATE: 2000-05-25
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 13
LENGTH: 51
TYPE: PRT
ORGANISM: Bordetella bronchiseptica
US-09-855-754B-13

Query Match 100.0%; Score 31; DB 5; Length 51;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFGP 5
|||||
DB 23 GGFGP 27

RESULT 4
US-09-855-754B-9
Sequence 9, Application US/09855754B
GENERAL INFORMATION:
APPLICANT: BOURSAX-EDDE, CAROLINE
TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED
TITLE OF INVENTION: REGIONS OF PERTACTIN IN BORDETELLA PERTUSSIS,
TITLE OF INVENTION: BORDETELLA PARAPERTUSSIS, AND BORDETELLA
TITLE OF INVENTION: BRONCHISEPTICA, THEIR USE IN DIAGNOSTICS, AND IN
FILE REFERENCE: 03495-0206-00000
CURRENT APPLICATION NUMBER: US/09/855,754B
CURRENT FILING DATE: 2001-05-16
PRIOR APPLICATION NUMBER: 60/206,969
PRIOR FILING DATE: 2000-05-25
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 9
LENGTH: 56
TYPE: PRT
ORGANISM: Bordetella bronchiseptica
US-09-855-754B-9

Query Match 100.0%; Score 31; DB 5; Length 56;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFGP 5
|||||
DB 28 GGFGP 32

RESULT 5
US-09-855-754B-10
Sequence 10, Application US/09855754B
GENERAL INFORMATION:
APPLICANT: BOURSAX-EDDE, CAROLINE
TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED
TITLE OF INVENTION: REGIONS OF PERTACTIN IN BORDETELLA PERTUSSIS,
TITLE OF INVENTION: BORDETELLA PARAPERTUSSIS, AND BORDETELLA
TITLE OF INVENTION: BRONCHISEPTICA, THEIR USE IN DIAGNOSTICS, AND IN
FILE REFERENCE: 03495-0206-00000
CURRENT APPLICATION NUMBER: US/09/855,754B
CURRENT FILING DATE: 2001-05-16
PRIOR APPLICATION NUMBER: 60/206,969
PRIOR FILING DATE: 2000-05-25
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 10
LENGTH: 56
TYPE: PRT
ORGANISM: Bordetella bronchiseptica
US-09-855-754B-10

Query Match 100.0%; Score 31; DB 5; Length 56;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFGP 5
|||||
DB 28 GGFGP 32

RESULT 6
US-09-855-754B-12
Sequence 12, Application US/09855754B
GENERAL INFORMATION:
APPLICANT: BOURSAX-EDDE, CAROLINE
TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED
TITLE OF INVENTION: REGIONS OF PERTACTIN IN BORDETELLA PERTUSSIS,
TITLE OF INVENTION: BORDETELLA PARAPERTUSSIS, AND BORDETELLA
TITLE OF INVENTION: BRONCHISEPTICA, THEIR USE IN DIAGNOSTICS, AND IN
FILE REFERENCE: 03495-0206-00000
CURRENT APPLICATION NUMBER: US/09/855,754B
CURRENT FILING DATE: 2001-05-16
PRIOR APPLICATION NUMBER: 60/206,969
PRIOR FILING DATE: 2000-05-25
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 12
LENGTH: 56
TYPE: PRT
ORGANISM: Bordetella bronchiseptica
US-09-855-754B-12

Query Match 100.0%; Score 31; DB 5; Length 56;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFGP 5
|||||


```
Db      23 GGFGP 27

RESULT 7
US-09-855-754B-11
; Sequence 11, Application US/09855754B
; GENERAL INFORMATION:
; APPLICANT: BOURSAX-ETUDE CAROLINE
; APPLICANT: GUISO-MACLOUF, NICOLE
; TITLE OF INVENTION: POLYPEPTIDES CONTAINING POLYMORPHISMS OF THE REPEATED
; TITLE OF INVENTION: REGIONS OF PERTACTIN IN BORDETELLA PERTUSSIS,
; TITLE OF INVENTION: BRONCHISEPTICA, AND BORDETELLA
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS
; FILE REFERENCE: 03495-0206-00000
; CURRENT APPLICATION NUMBER: US/09/855, 754B
; CURRENT FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 60/206, 969
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 61
; TYPE: PRT
; ORGANISM: Bordetella bronchiseptica
US-09-855-754B-11

Query Match
Best Local Similarity 100.0%; Score 31; DB 5; Length 61;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GGFGP 5
       11111
Db      23 GGFGP 27

RESULT 8
US-09-724-676-86569
; Sequence 86569, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724, 676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 86569
; LENGTH: 69
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-86569

Query Match
Best Local Similarity 100.0%; Score 31; DB 5; Length 69;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GGFGP 5
       11111
Db      47 GGFGP 51

RESULT 9
US-09-724-676A-86569
; Sequence 86569, Application US/09724676A
; GENERAL INFORMATION: LTD
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724, 676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
US-09-724-676A-86569

; SEQ ID NO 86569
; LENGTH: 69
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-86569

Query Match
Best Local Similarity 100.0%; Score 31; DB 5; Length 69;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GGFGP 5
       11111
Db      47 GGFGP 51

RESULT 10
PCT-US02-32727-7133
; Sequence 7133, Application PC/TUS0232727
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Skelky, Yasir
; APPLICANT: Persing, David
; APPLICANT: Bhatia, Ajay
; APPLICANT: Maisonneuve, Jean Francois
; APPLICANT: Zhang, Yanni
; APPLICANT: Wang, Siding
; APPLICANT: Jen, Shyian
; APPLICANT: Lodes, Michael
; APPLICANT: Benson, Darin
; APPLICANT: Jones, Robert
; APPLICANT: Carter, Darriack
; APPLICANT: Barth, Brenda
; APPLICANT: Douglass, John
; FILE REFERENCE: 210121.514C1
; CURRENT APPLICATION NUMBER: PCT/US02/32727
; CURRENT FILING DATE: 2002-10-11
; NUMBER OF SEQ ID NOS: 30992
; SEQ ID NO 7133
; LENGTH: 74
; TYPE: PRT
; ORGANISM: Propionl acnes
PCT-US02-32727-7133

Query Match
Best Local Similarity 100.0%; Score 31; DB 1; Length 74;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GGFGP 5
       11111
Db      54 GGFGP 58

RESULT 11
US-09-978-825-7133
; Sequence 7133, Application US/09978825
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Skelky, Yasir
; APPLICANT: Persing, David
; APPLICANT: Bhatia, Ajay
; APPLICANT: Maisonneuve, Jean Francois
; APPLICANT: Zhang, Yanni
; APPLICANT: Wang, Siding
; APPLICANT: Jen, Shyian
; APPLICANT: Lodes, Michael
; APPLICANT: Benson, Darin
; APPLICANT: Jones, Robert
; APPLICANT: Carter, Darriack
; APPLICANT: Barth, Brenda
; APPLICANT: Douglass, John
; FILE REFERENCE: 210121.514C1
US-09-978-825-7133
```

CURRENT APPLICATION NUMBER: US/09/978,825
CURRENT FILING DATE: 2003-01-29
NUMBER OF SEQ ID NOS: 30992
SEQ ID NO 7133
LENGTH: 74
TYPE: PRT
ORGANISM: Propionl acnes
US-09-978-825-7133

Query Match 100.0%; Score 31; DB 5; Length 74;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFGP 5
Db 54 GGFGP 58

RESULT 12
US-10-057-498-7133
GENERAL INFORMATION:
APPLICANT: Mitcham, Jennifer
APPLICANT: Skeiky, Yasir
APPLICANT: Persing, David
TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acnes
FILE REFERENCE: 210121.514
CURRENT APPLICATION NUMBER: US/10/057,498
CURRENT FILING DATE: 2001-04-20
NUMBER OF SEQ ID NOS: 29212
SEQ ID NO 7133
LENGTH: 74
TYPE: PRT
ORGANISM: Propionl acnes
US-10-057-498-7133

Query Match 100.0%; Score 31; DB 6; Length 74;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFGP 5
Db 54 GGFGP 58

RESULT 13
PCT-US02-32727-4050
GENERAL INFORMATION:
APPLICANT: Mitcham, Jennifer
APPLICANT: Skeiky, Yasir
APPLICANT: Persing, David
APPLICANT: Bhatia, Ajay
APPLICANT: Maisonneuve, Jean Francois
APPLICANT: Zhang, Yannl
APPLICANT: Wang, Siqing
APPLICANT: Jen, Shylan
APPLICANT: Lodes, Michael
APPLICANT: Benson, Darin
APPLICANT: Jones, Robert
APPLICANT: Carter, Darrick
APPLICANT: Barth, Brenda
TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acnes
FILE REFERENCE: 210121.514C1
CURRENT APPLICATION NUMBER: PCT/US02/32727
CURRENT FILING DATE: 2002-10-11
NUMBER OF SEQ ID NOS: 30992
SEQ ID NO 4050
LENGTH: 80
TYPE: PRT
ORGANISM: Propionl acnes
PCT-US02-32727-4050

Query Match 100.0%; Score 31; DB 1; Length 80;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFGP 5
Db 57 GGFGP 61

RESULT 14
US-09-978-825-4050
GENERAL INFORMATION:
APPLICANT: Mitcham, Jennifer
APPLICANT: Skeiky, Yasir
APPLICANT: Persing, David
APPLICANT: Bhatia, Ajay
APPLICANT: Maisonneuve, Jean Francois
APPLICANT: Zhang, Yannl
APPLICANT: Wang, Siqing
APPLICANT: Jen, Shylan
APPLICANT: Lodes, Michael
APPLICANT: Benson, Darin
APPLICANT: Jones, Robert
APPLICANT: Carter, Darrick
APPLICANT: Barth, Brenda
TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne
FILE REFERENCE: 210121.514C1
CURRENT APPLICATION NUMBER: US/09/978,825
CURRENT FILING DATE: 2003-01-29
NUMBER OF SEQ ID NOS: 30992
SEQ ID NO 4050
LENGTH: 80
TYPE: PRT
ORGANISM: Propionl acnes
US-09-978-825-4050

Query Match 100.0%; Score 31; DB 5; Length 80;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFGP 5
Db 57 GGFGP 61

RESULT 15
US-10-057-498-4050
GENERAL INFORMATION:
APPLICANT: Mitcham, Jennifer
APPLICANT: Skeiky, Yasir
APPLICANT: Persing, David
TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne
FILE REFERENCE: 210121.514
CURRENT APPLICATION NUMBER: US/10/057,498
CURRENT FILING DATE: 2001-04-20
NUMBER OF SEQ ID NOS: 29212
SEQ ID NO 4050
LENGTH: 80
TYPE: PRT
ORGANISM: Propionl acnes
US-10-057-498-4050

Query Match 100.0%; Score 31; DB 6; Length 80;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGFGP 5
Db 57 GGFGP 61

Mon Mar 17 13:01:38 2003

seq25-fg.rpn

Page 5

Search completed: March 17, 2003, 12:20:53
Job time : 24 secs

